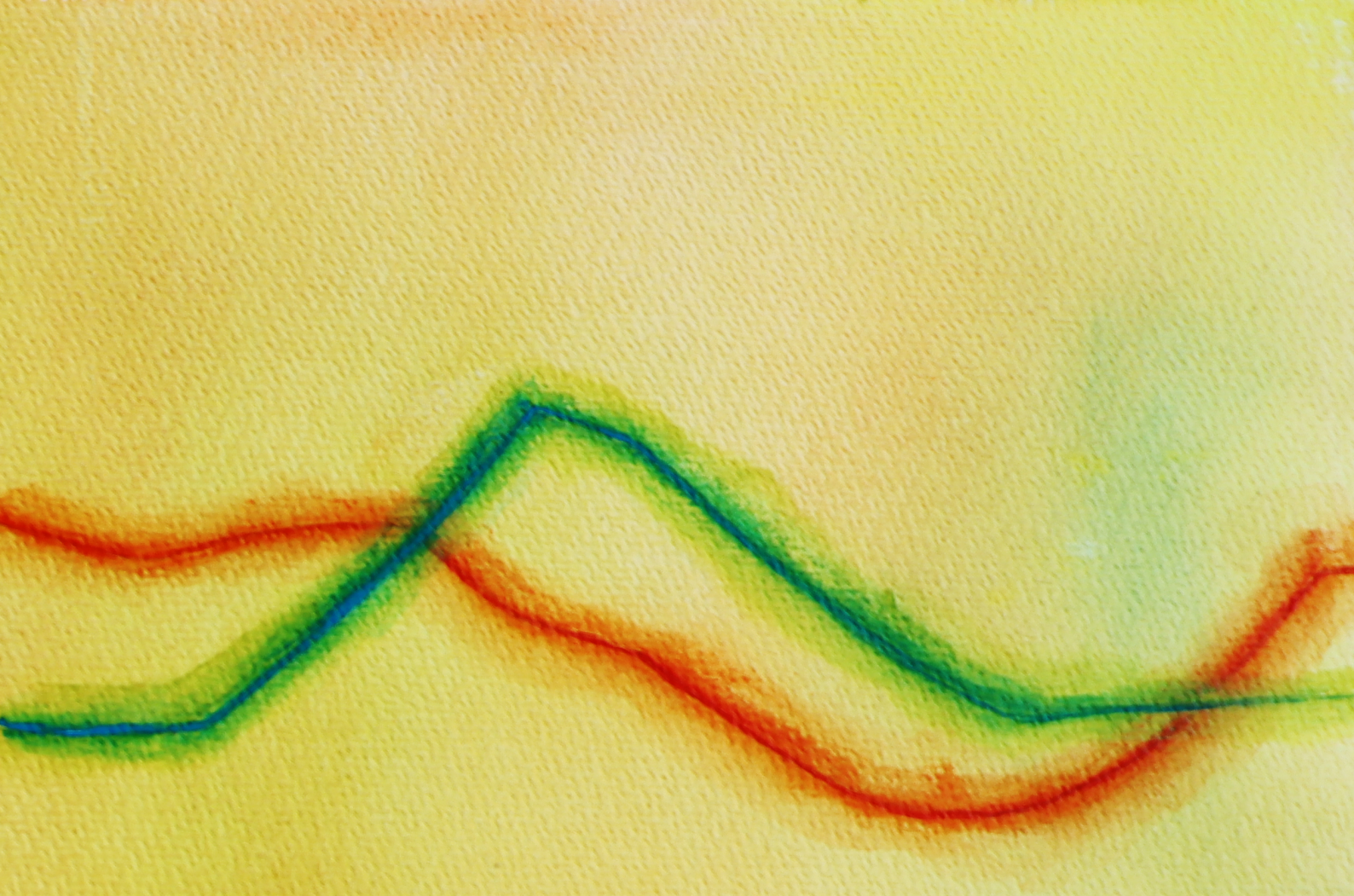


# Climate Change Impacts on Hydrology

## Informing the Design of the Modeling Chain and Supporting Adaptation

Kirsti Hakala

# Climate Change Impacts on Hydrology

## Informing the Design of the Modeling Chain and Supporting Adaptation

Dissertation
zur
Erlangung der naturwissenschaftlichen Doktorwürde
(Dr. sc. nat.)

vorgelegt der
Mathematisch-naturwissenschaftlichen Fakultät
der
Universität Zürich
von

**Kirsti Hakala**
aus den
Vereinigten Staaten von Amerika

Promotionskommission
**Prof. Dr. Jan Seibert (Vorsitz)**
**Dr. Nans Addor**
**Prof. Dr. Reinhard Furrer**

Zürich, 2020

Kirsti Hakala

Climate Change Impacts on Hydrology -
Informing the Design of the Modeling Chain and Supporting Adaptation

PhD thesis, University of Zurich, Zurich, Switzerland (2020)

DOI: https://doi.org/10.5167/uzh-199631

*For my mother...*

# Abstract

As climate change continues to perturbate regional hydrology, it has become increasingly necessary for the scientific community to support water management decisions. This responsibility requires an understanding of the impact of climatic and environmental change on future hydrological conditions. Numerical models are the principal tools used to obtain quantitative information on these dynamics. In common practice, multiple models are used in succession, known as a modeling chain. To assess and simulate potential changes to a hydrological regime due to climate change, it is first necessary to acquire climate information for current and future conditions. Emission scenarios are utilized to extend current knowledge of historical greenhouse gas emission data into the future. These scenarios are then used as input to general circulation models (GCMs) to provide information about how Earth system processes are likely going to develop under the influence of climate change. However, the spatial scale of GCMs is too coarse for local-scale impact studies. Thus, regional climate models (RCMs) are used to dynamically downscale climate variables. However, even after downscaling, biases often still exist between the GCM-RCM output and local-scale observational data. Bias correction techniques are therefore employed to reduce these biases, resulting in climate variables that are suitable as input to hydrological models. Hydrological models are then calibrated and used to create projections of streamflow under the influence of climate change.

This dissertation delves into the modeling approach to achieve projections of climate change impacts on water resources. An initial study reviewed and summarized the steps to perform a climate change impact analysis as a chapter contribution to the Encyclopedia of Water. The motivation of this encyclopedia chapter was to provide guidance on the modeling chain, which was previously lacking in the literature. Another study introduced a process-based approach to evaluate the performance of climate models and their bias correction. This new approach was based on the concept that climate change impacts are the result of the interactions between variables, which are not considered in typical analyses. Additionally, the sensitivity of the modeling chain was evaluated in two co-authored papers, which showed that discharge projections are not especially sensitive to varying levels of potential evapotranspiration information in arid regions. However, discharge projections were shown to be sensitive to the choice of bias correction method for two glaciated catchments within Switzerland. The choice of bias correction method translated into considerable consequences for the hydrological responses of the catchments, although differences in total streamflow were negligible. The lessons learned from these aforementioned studies were applied in a final study, which aimed to support the decision

making of a Swiss hydroelectricity company. Following a user-centered approach, a tailored climate change impact analysis was carried out. This approach allowed for insightful recommendations for climate change adaptation and decision-support regarding their upcoming concession negotiations. This study highlighted the importance of accounting for a stakeholder's specific needs when designing a climate change impact study.

**Key words:** Climate Change, Bias Correction, Decision Support, Hydrological Modeling, HBV, Climate Models, Ensembles, GCM, RCM, Modeling Chain, Streamflow, Potential Evapotranspiration

# Contents

# Acronyms

## Modeling chain

| | |
|---|---|
| RCP | Representative concentration pathway |
| GHG | Greenhouse gas |
| GCM | General circulation model |
| RCM | Regional climate model |
| GCM-RCM | Regional climate model forced by a general circulation model |
| DD | Dynamical downscaling |
| SD | Statistical downscaling |
| *QM/QDM | Univariate quantile mapping |
| MBCn | Multivariate bias correction |
| *Raw/noBC | Uncorrected data |
| GAP | Genetic Algorithm and Powell optimization |
| CDF | Cumulative distribution function |
| ANOVA | Analysis of variance |

## Variables

| | |
|---|---|
| T | Temperature |
| P | Precipitation |
| $^{**}E_{pot}$ | Potential evapotranspiration |
| $^{**}E_{act}$ | Actual evapotranspiration |
| SWE | Snow water equivalent |
| V | Glacier ice volume |
| $R_e$ | Extraterrestrial radiation |
| $\lambda$ | Latent heat flux |
| $\rho$ | Density of water |

## Hydrological models

| | |
|---|---|
| HBV | Hydrologiska Byråns Vattenbalansavdelning |
| GR4J | Modèle du Génie Rural à 4 paramètres Journaliers |
| IHACRES | Identification of unit Hydrograph and Component flows from Rainfall, Evapotranspiration and Streamflow |

## Entities and projects

| | |
|---|---|
| EURO-CORDEX | European Coordinated Regional Downscaling Experiment |
| CMIP5 | Coupled Model Intercomparison Project- Phase 5 |
| EU | European Union |
| UNEP | United Nations Environment Programme |
| ECMWF | European Centre for Medium-Range Weather Forecasts |
| ERA-Interim | ECMWF re-analysis product covering 1979- present |
| BoM | Australian Bureau of Meteorology |
| SMHI | Swedish Meteorological and Hydrological Institute |

## Objective functions

| | |
|---|---|
| $R_{eff}$ | Nash-Sutcliffe efficiency |
| $R_{eff\_log}$ | Nash-Sutcliffe efficiency log |
| $R_{eff\_JS}$ | Nash-Sutcliffe efficiency for June September |
| $R_{KGE}$ | Kling-Gupta efficiency |
| $R_{Lindström}$ | Lindström measure |
| $R_{VE}$ | Volume error |
| RMSE | Root mean square error |
| $MAE_{Normalized}$ | Mean absolute normalized error |

## Time series

| | |
|---|---|
| $Q_{obs}$ | Observed discharge |
| $Q_{ref}$ | Simulated discharge (hydrological model is forced by observed climate data) |
| $Q_{raw}$ | Simulated discharge (hydrological model is forced by raw GCM-RCM data) |
| $Q_{qm}$ | Simulated discharge (hydrological model is forced by bias-corrected GCM-RCM data) |
| $R_{qm}$ | Contribution of rain to streamflow (hydrological model is forced by bias-corrected GCM-RCM data) |
| $S_{qm}$ | Streamflow resulting from snowfall (hydrological model is forced by bias-corrected GCM-RCM data) |
| $T_{obs}$ | Observed temperature |
| $T_{raw}$ | Simulated temperature (stemming from raw GCM-RCM data) |
| $T_{qm}$ | Simulated temperature (stemming from bias-corrected GCM-RCM data) |
| $P_{obs}$ | Observed precipitation |
| $P_{raw}$ | Simulated precipitation (stemming from raw GCM-RCM data) |
| $P_{qm}$ | Simulated precipitation (stemming from bias-corrected GCM-RCM data) |

## Indices

| | |
|---|---|
| Q5 | 5th percentile of daily streamflow |
| Q95 | 95th percentile of daily streamflow |
| $H_{DD}$ | Heating degree days |
| $C_{DD}$ | Cooling degree days |

## Streamflow components

| | |
|---|---|
| $Q_R$ | Streamflow resulting from rainfall |
| $Q_S$ | Streamflow resulting from snowfall |
| $Q_I$ | Streamflow resulting from icemelt |

## Seasons

| | |
|---|---|
| DJF | December, January, February |
| MAM | March, April, May |
| JJA | June, July, August |
| SON | September, October, November |

## Other climate models

| | |
|---|---|
| EBM | Energy balance model |
| RC | Radiative convective model |
| SDM | Statistical dynamic model |

**Both acronyms are used throughout the dissertation in order to stay consistent with the term used within the paper being discussed. **$E_{pot}$ and $E_{act}$ are synonymous with PET and AET respectively within Paper IV.*

# Papers and Author Contributions

## List of papers

I Hakala, K., N. Addor, C. Teutschbein, M. Vis, H. Dakhlaoui, and J. Seibert (2019), Hydrological modeling of climate change impacts. Encyclopedia of Water: Science, Technology, and Society, 1-20, doi:10.1002/9781119300762.wsts0062.

II Hakala, K., N. Addor, and J. Seibert (2018), Hydrological modeling to evaluate climate model simulations and their bias correction. Journal of Hydrometeorology, 19, 1321-1337, doi:10.1175/JHM-D-17-0189.1.

III Meyer, J., I. Kohn, K. Stahl, K. Hakala, J. Seibert, and A. J. Cannon (2019), Effects of univariate and multivariate bias correction on hydrological impact projections in alpine catchments. Hydrological Earth System Sciences, 23, 1339-1354, doi:10.5194/hess-23-1339-2019.

IV Dakhlaoui, H., J. Seibert, and K. Hakala (2020), Sensitivity of discharge projections to potential evapotranspiration estimation in Northern Tunisia. Regional Environmental Change, 20, 1-12, doi:10.1007/s10113-020-01615-8.

V Hakala, K., N. Addor, T. Gobbe, J. Ruffieux, J. Seibert (2020), Risks and opportunities for a Swiss hydroelectricity company in a changing climate. Hydrology and Earth System Sciences, 24, 3815-3833, doi:10.5194/hess-24-3815-2020.

## Author Contributions

Paper I: As an invited paper, the scope of the paper was discussed initially with all authors, and the outline was mutually agreed upon. Designated authors were chosen for specific subsections, with contributions received as follows: Nans Addor- uncertainty decomposition, sampling modelling space, Claudia Teutschbein- downscaling and bias correction, Marc Vis- hydrological model input and calibration/validation, Hammouda Dahklaoui- methods to assess parameter uncertainty, Jan Seibert- hydrological model complexity and parameter instability, and myself- I contributed extensively to all sections and took the lead on developing a cohesive first draft. Further manuscript development primarily occurred between myself and Nans Addor. All authors contributed to finalizing the manuscript.

Paper II: Jan Seibert conceived the initial ideas for this study. I performed the simulations, analyzed the results, and wrote the first draft of the manuscript with iterative feedback from Nans Addor. All authors contributed to the interpretation of final results and manuscript development.

Paper III: Judith Meyer, Irene Kohn, Kerstin Stahl, and Jan Seibert designed the study. Judith Meyer carried out bias correction, modeling, and analyses and wrote the first draft of the manuscript. Irene Kohn calibrated the hydrological model and prepared snow, glacier, and hydrological data. I prepared the EUROCORDEX data for the catchments. Alex Cannon provided and helped with his bias correction scripts. All authors contributed to and edited the paper.

Paper IV: Jan Seibert and Hamouda Dahklaoui conceived the initial ideas for this study. I prepared the EUROCORDEX data for the catchments. Hamouda Dahklaoui performed the simulations and analyses. Results were discussed amongst all authors. Hamouda Dahklaoui took the lead on developing the first draft of the manuscript. The manuscript was further developed primarily between myself and Hamouda Dahkaloui.

Paper V: Nans Addor and I designed the study based on previous exchanges between Nans Addor and Groupe E. I refined the scope of the project with Groupe E (Johann Ruffieux and Thibault Gobbe) over the course of several meetings. I performed the climate change impact analysis and discussed the results with all coauthors. Writing of this paper was led by myself with contribution from all coauthors. Finalization of the manuscript occurred primarily between myself and Nans Addor.

# Chapter 1

# Introduction

## 1.1 Setting the stage for climate change studies

On a cold fall day in October 2019, I found myself seated in a large auditorium at the World Meteorological Organization in Genève. Scientists and water managers from all over the world had convened for the High Mountain Water Summit. The goal of the summit was to develop a call to action, where the scientific community could agree and work together on common issues facing us all. Again and again, every panelist and scientific leader who spoke recognized that one of the most critical manifestations of anthropogenic climate change is its effect on the water cycle, which is echoed throughout the literature (e.g. Alexander et al., 2006; Huntington, 2006; IPCC, 2014; Hirabayashi et al., 2013; Arora and Boer, 2001; Stahl et al., 2010). Summit speakers vehemently spoke about the struggles within their own country and common concerns were identified. One of the most widely mentioned concerns was that pressures from the amplification of anthropogenic climate change are causing an unprecedented crisis that threatens the sustainability of the planet. Leaders expressed concern that changes to water availability have important ramifications for ecosystems, the cryosphere, and the capacity for these environments to support livelihoods and local economies and industries (e.g. hydropower, agriculture, transportation, tourism). In short, there is a demand and a movement towards combating climate change and its impact on water resources.

In order to help address these issues, hydrological modelers have been tasked with providing the analytical basis for the choices and investments made by policy makers and water managers. Numerical models are appropriate tools to provide quantitative information on the development of the environment under the influence of climate change. In practice, multiple models are used to achieve projections that are useful for decision makers. This multi-model method is often referred to as a modeling chain within climate change literature. This dissertation delves into the numerical modeling approach to achieve projections of climate change impacts on water resources. Over the course of five dissertation papers, we (1) review the most relevant subcomponents of hydrological climate change research and provide guidance on the modeling chain, (2) rethink the evaluation of the modeling chain, (3) alter particular aspects of the modeling chain to evaluate sensitivity, and (4) apply a modeling chain to a real-world decision-making problem.

## 1.2 Overview of modeling chain

The end-goal of a typical modeling chain is to produce projections of streamflow under the influence of different climate change scenarios. Streamflow within a catchment is the result of processes occurring at a wide range of spatial scales. A modeling chain inherently comprises of multiple models, where each model incorporates the physics and processes of a certain scale. This is based on the premise that processes are best represented at the scale at which they occur. Therefore, models are combined in succession to utilize their best strengths of simulating processes they were designed for. In an iterative manner, a modeling chain progressively translates global-scale information to finer scales. The individual steps of a typical modeling chain are shown in Figure 1.1, and are briefly summarized in the following sections. A detailed description of the modeling chain is provided within Paper I of this dissertation. Throughout the text, I use the pronoun 'we' when referring to any efforts related to the dissertation papers as these studies were always the result of collaborations.

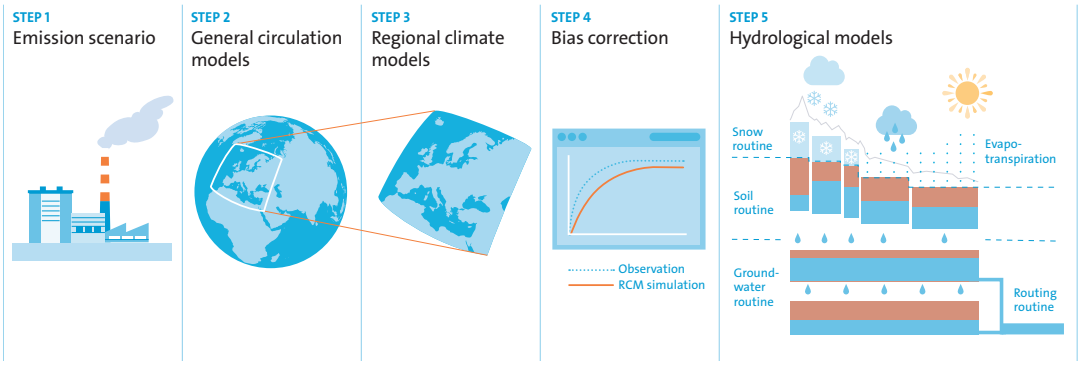



FIGURE 1.1: Depiction of a typical modeling chain, used to create projections of streamflow under the influence of climate change. Figure adapted from Paper I.

## 1.3 Emission scenarios

How will the world's climate change in the future? The answer to this question is dependent on how human societies choose to develop. Emission scenarios utilize historical greenhouse gas (GHG) emission data and extend this knowledge into the future using estimates of how emissions may progress with evolving demography, economic development, technology, as well as energy generation and land use change. Emission scenarios are the brainchild of the Intergovernmental Panel on Climate Change (IPCC), established in 1988 by two United Nations organizations, the World Meteorological Organization (WMO) and the United Nations Environment Programme (UNEP). At the time of this dissertation, the IPCC has published five assessment reports. In its Fifth Assessment Report (AR5), the latest set of emission scenarios were employed, namely the Representative Concentration Pathways (RCPs). The RCPs are based on four selected scenarios produced from four modeling teams. Unlike their predecessor, the SRES, the RCPs are not a complete package

of socioeconomic, emission, and land use change projections. The IPCC decided to discontinue the SRES because they were seen as too confining given that certain projected changes were inherently bundled with one another (i.e. SRES can be thought of as 'what if' situations). Rather, the RCPs are described as 'pathways' given that they relate to concentrations of GHGs in the atmosphere and their cumulative budgets. The RCPs are therefore rather end-points which can be reached via multiple pathways. In general, the RCPs represent different rates and magnitudes of climate change, which can be used to assess the risk of crossing certain thresholds of change. In practical terms, the RCPs consist of numerical data, and are stored in databases. Modelers can then download these database sets and and use them to force their models in order to produce projections of the environment under the influence of climate change.

## 1.4 General circulation models

General circulation models (GCMs) are tools which embody simplified mathematical representations of the climate system (IPCC, 2001). They are used to aid in the understanding of the climate system by simulating and predicting the behaviour of different variables under a future climate. Within a modeling chain, GCMs are run with emission scenarios as forcing to create projections of the climate under changing GHG concentrations. GCMs are based on the principles of fluid dynamics and thermodynamics, where major process such as the circulation of the ocean and atmosphere are simulated. We know from climate reconstructions based on ice cores, sediment cores, tree rings, etc. that the Earth's climate has experienced oscillations over time. The data show that the Earth's climate occurs in cycles with predictable periods. GCMs are a state-of-the-art tool in that they are designed to simulate the mechanisms leading to these oscillations, allowing modelers to investigate and understand the behaviour of past, present and future climate.

Climate modeling has been around for over fifty years with the advent of some of the first global atmospheric models in the 1960s (e.g., Manabe and Wetherald, 1967). Over the years, numerous climate models have become available which range in complexity from one dimensional to three-dimensional climate models. Climate models can generally be classified based on their dimensionality. Mcguffie and Henderson-Sellers (2001) have classified climate models into the following categories, (1) zero or one-dimensional energy balance models (EBMs) that predict Earths surface temperature by balancing incoming and outgoing radiation of the planet (if 1-D then latitudinal energy transfer is also accounted for), (2) one or two-dimensional radiative convective (RC) models that model the temperature profile of the atmosphere by considering the transport of energy through the depth of the atmosphere (if 2-D then horizontally-averaged energy transfers are also simulated), (3) two-dimensional, statistical-dynamical models (SDMs) that model temperature within a 2-D framework (e.g. latitude and height) and can be thought of as a combination of EBMs and RC models, and (4) GCMs which allow for variation in a 3-D

framework (latitude, longitude, vertical 'height'), where the Earth can divided into domains with each domain having its own set of values for particular climate parameters.

Although the climate models described above evolved essentially concurrently, by the 1980s and 1990s the movement to channel funding and efforts towards the development of GCMs became overwhelming (Mcguffie and Henderson-Sellers, 2001). While the simpler models have been used to provide globally- or regionally-averaged estimates of climate change, GCMs are the only climate models which have the ability to provide both spatially and physically realistic estimates of climate change, which is one of the first requirements for any impact analysis. The spatial resolution of GCMs has progressively increased with each new generation of models. Currently GCMs operate at horizontal resolutions ranging between 110-280 km and 10 to 30 levels in the vertical. For the sake of most water management decisions, the GCMs spatial scale is insufficient, because it is lacking detailed regional information (IPCC, 2007). Thus, downscaling procedures are required in order to derive high-resolution climate parameters for hydrological modeling.

## 1.5 Downscaling

Downscaling can be defined as the transfer of large-scale information to a smaller-scale, resulting in an increase in resolution of the data. To transfer coarse-resolution GCM output to a spatial resolution appropriate for hydrological modeling, numerous downscaling techniques have emerged. In general, these techniques can be classified into either statistical (SD) or dynamical downscaling (DD). SD has its roots within weather forecasting, where it was first used to establish statistical relationships between large-scale and regional-scale variables (Maraun and Widmann, 2018). Given the interconnectedness between SD and bias correction, the following subsection on bias correction describes this history in greater detail. SD has the benefit of being computationally cheap, but is also limited in that it is based on statistical relationships that are assumed to not change over time.

DD implies the use of higher-resolution (10-50km) regional climate models (RCMs) for limited regions with boundary conditions based on large-scale information provided by GCMs. This relationship between GCMs and RCM has also been coined as the 'nested' approach. DD was first applied in a climate change study in the late 1980's by Dickinson et al. (1989). DD has the benefit of producing internally consistent output variables, however, it has some limitations due to commensurability issues and simplification of complex processes. These simplifications can cause systematic model errors originating either from inadequately defined model structures or parameters. Thus, there is the need to further post-process RCM data using bias correction.

## 1.6 Bias correction

Many bias correction techniques have been developed to adjust climate model simulations in order to reduce the effects of systematic errors. Bias correction methods range in their degree of complexity (i.e., how many statistical moments they are able to correct). The spectrum of bias correction methods range from simple-to-apply methods such as scaling factors to more sophisticated methods such as distribution mapping. Among the different bias correction methods, distribution mapping has been identified as the most efficient one in adjusting RCM simulations (see Teutschbein and Seibert (2012) or Chen et al. (2013) for a review). It tries to match the distribution of the RCM-simulated climate values with the observed distribution with the help of a transfer function. Typically, bias correction methods are one-dimensional or univariate, i.e., they adjust only one RCM-simulated variable at a time without assuring consistency between the climate variables (Maraun, 2016). However, climate change impacts are the result of intervariable relationships, which are not considered in such methods. In Paper II, we introduce a more process-orientated method for analyzing climate model and bias correction performance, which accounts for intervariable relationships within its design. This point is also addressed in Paper III, where we show that the use of a multivariate versus a univariate bias correction method can lead to different conclusions drawn for an impact study.

Bias correction has an intertwined history with SD to the point that some research papers will use the terms SD and bias correction interchangeably. However, there are distinct differences in their origins. SD was first implemented in the 1940's, whereas bias correction came about decades later. In its origin, SD was primarily used to downscale weather forecasts. Specifically it was used to infer a statistical relationship between large-scale observational information (predictor) and an observed local-scale variable (predictand). This statistical relationship was then applied to downscale a large-scale weather forecast to the local-scale. However, in this context, the modeled forecasts were assumed to be perfectly fitted with the large-scale observations, known as perfect prognosis. However, it was noted that even after the application of SD, the forecasts were still significantly deviated from observation. This led to the use of model output statistics (MOS), which infers a statistical relationship between a large-scale modeled predictor and an observed local-scale predictand. Model biases are therefore already accounted for during MOS, which is distinctly different compared to SD. The use of the term 'bias correction' varies greatly between different communities. The concept of bias correction can be thought of as a special case of MOS. A brief summary of this history is provided in Paper I. A detailed discussion of the origins of bias correction and SD is available within Chapters 3 and 12 within the book by Maraun and Widmann (2018). Once the RCM climate data has been bias corrected, it is then ready to be used as input to a hydrological model.

## 1.7 Estimation of potential evapotranspiration

The practice of hydrological modeling often utilizes the evaporation of water in its calculations of runoff production and thus runoff discharge. Changes to future evaporative demands are likely to be as important as changes to precipitation patterns in determining the behavior of river flows (Prudhomme and Williamson, 2013). Since different research communities often define evapotranspiration and potential evaporation using various nomenclatures, it is important to state definitions directly. Evaporation here can be considered an overarching term to encompass all processes in which liquid water is transferred as water vapor to the atmosphere (McMahon et al., 2016). Evapotranspiration is defined as the sum of two fluxes: transpiration from plants and soil evaporation. Evapotranspiration can be said to occur at a potential rate ($E_{pot}$) when availability of water is not limited (Bartholomeus et al., 2015). In order to differentiate between evapotranspiration at its potential rate ($E_{pot}$) or actual/real rate, the term actual evapotranspiration ($E_{act}$) is used throughout this dissertation.

The modern definition of $E_{pot}$ modeling dates back to Thornthwaite and slightly later Penman who both in 1948 put forth the concept of $E_{pot}$ as a tool for modeling evaporation for a short green crop completely shading the ground (McMahon et al., 2016). In semi-arid environments there exists a large excess of $E_{pot}$ over rainfall, making $E_{pot}$ an important concept for irrigators in those regions. Hydrologists find $E_{pot}$ equally important, as it can be a useful tool for estimating streamflow under a variety of climatic conditions. Since 1948, many different $E_{pot}$ methods have been developed. $E_{pot}$ methods are a highly debated topic within the scientific literature and many publications exist for the purpose of comparing them (e.g. Oudin et al., 2005; Prudhomme and Williamson, 2013; Thompson et al., 2014; Bartholomeus et al., 2015). Each method has its own data requirements, efficiency and designed purpose.

Dry lands (semi arid or arid regions) cover approximately 40% of the Earth's land surface and by the year 2030, approximately half of the world's population will be living in areas of high water stress (UNEP, 2015). World-wide two thirds of terrestrial precipitation is lost through evaporation (McMahon et al., 2016), however in arid regions such as the Mediterranean region, this proportion is closer to 90%. Evaporation is an essential part of the water balance at the catchment scale, so it may be wondered at this stage why one should concern themselves with evaporation at a hypothetical rate ($E_{pot}$). However, compared to streamflow and precipitation, $E_{pot}$ is more difficult to measure directly in the field. Since undisturbed natural plant and soil conditions cannot be properly reproduced instrumentally, there is always some doubt about the measurements of $E_{act}$. Whereas, $E_{pot}$ can be calculated using climate data and can be simulated with less advanced instruments. In this way, $E_{pot}$ is relatively easier to estimate and is often used as input for hydrological modeling. Methods for estimating $E_{pot}$ also vary in their complexity, however data availability is often a limiting factor preventing the use of physically-based equations.

## 1.8 Hydrological models

At this stage within the modeling chain climate data has now been prepared as input to force a hydrological model(s), which then in turn is used to generate projections of streamflow. There are a range of hydrological models to choose from, which vary largely in their process complexity and their spatial resolution. The question of which model to use and what complexity is necessary (both spatial and physical), is not an easy thing to answer. Hydrological models should be selected based on their intended purpose so it is necessary to first consider the requirements of the project and the desired output (Ková et al., 2015; Lee, 1993). The 1960's saw the advent of computer-based models used to simulate hydrological regimes, where individual components of the hydrological cycle were represented by conceptual elements (Todini, 2007; Todini, 2011). Over time, a plethora of models were developed, each representing the responses and interactions of these hydrological components in different ways. Fast forward to today, and you will find a vast selection of hydrological models available for the purpose of climate change impact studies. As Hrachowitz and Clark (2017) note, a continuum exists along which hydrological models can be described. At one end of the spectrum, there exists the lumped, conceptual models with 5-20 model parameters, which treat the catchment as one lumped unit (although usually considering different elevation zones). At the other end of the spectrum lies fully distributed, physically-based models which use hundreds of model parameters. Hydrological model taxonomy follows a loose naming framework. Model descriptions are often known by many synonymous names (e.g. distributed, physically-based, process-based). Hydrological models follow many colloquial naming systems and there is the tendency for models to fall within numerous categories (i.e. continuum of model complexity). Hrachowitz and Clark (2017) offers a thorough review of the pros and cons of different model complexities.

While physically-based models should be superior, in practice simpler models (i.e. conceptual models) often perform at least as good as the more complex models with regard to catchment runoff (e.g., Breuer et al., 2009). The pragmatic use of conceptual models is described well by one the early developers, Sten Bergström. He stated in 1992, that his work followed, 'to a high degree, the ideas lined out by Nash and Sutcliffe (1970) and Krause and Boyle (2005), among others, who saw the risk that increasing computer capacities may result in too complex model formulations, unless the significance of model components is carefully checked.' This philosophy is responsible for the relative simplicity of some hydrological models. As Bergström points out, without observational data to validate added complexity to a model, a simplistic structure should be chosen first to allow for an efficient model.

Hydrological models consist of parameters that are used to represent the storage and fluxes of water within a catchment. Equations in a conceptual model usually do not represent a singular physical processes, and therefore parameters cannot be linked to a specific property in the field. And even if they could be linked, they often represent average values

over larger areas (e.g. on catchment scale). As a result, parameter values can often not be measured directly. Instead, calibration techniques can be used to find optimal parameter values. Calibration entails using observational meteorological and discharge data as input into the hydrological model. The model is then calibrated, whereby the difference between simulated and observed discharge is minimized according to an objective criteria (i.e. objective function), resulting in a best parameter set. Several objective functions are available to evaluate the resemblance between observed and simulated discharge (e.g. Nash-Sutcliffe Efficiency, Kling-Gupta Efficiency), which each focus on certain aspects of the hydrograph. Calibration can be performed manually (i.e. one adjusts the parameters via a trial and error procedure to see which combination of parameter values results in a high objective function score). However, manual calibration can be tedious and time consuming. Consequently, automatic calibration methods have developed over time. Monte Carlo methods are some of the most commonly used automatic approaches for model calibration. Monte Carlo methods rely on repeated random sampling to obtain best performing parameter sets. Other more sophisticated methods include genetic algorithms, which apply parameter selection in a way resembling natural selection in biology (Wang, 1991). With the genetic algorithm, a population of $m$ points are chosen initially at random in the parameter space. The objective functions are then calculated at all points and then compared. From these points, two points are selected randomly with better points given higher chances for selection, with an occasional random distance added to the selection process (as specified by the modeler). This step is then repeated until $m$ new points are generated. The newly generated population is expected to be more closely concentrated around the optima compared to the original points. The new population of points can then be used to generate another population and so on. By repeating this entire process a multitude of times, it is gradually more likely that the final selection of points will be located around the vicinity of the global optimum, where the best parameter set should theoretically be found. Once a best set (or multiple best sets) of parameters are chosen, it is important to verify that the parameters achieved are reasonable. Verification can be done by validating the model for a situation that is different than that which it was calibrated on. This is an important step within climate change impact studies, where the future climate is expected to be different than the historical climate. To this extent, validation checks whether the selected parameter sets are reasonable under both a historical and future climate. An example of a validation method that is often used is the split-sample test, where the available data is split into sub-periods (e.g. two equal sub-periods). One sub-period can be used for calibration, and and another can be used for validation (and vice versa). Depending on the exact scenario, sub-periods can simply be chosen randomly or, for example, selected based on differences in climatic conditions. Acceptable validation results are a minimum requirement for having confidence in a model and its parameter values.

## 1.9 Characterizing and accounting for uncertainty

One of the greatest challenges related to climate change impact studies is the considerable amount of uncertainty revealed during the use of a modeling chain. Uncertainties stem from a myriad of sources, including the associated emission scenario, climate model members, downscaling techniques, model structures, initialization settings, boundary conditions, bias correction methods, hydrological models, model parameters, calibration periods, and the list certainly goes on. The uncertainties from one step of the modeling chain are passed on to the following steps in a cascading manner (Wilby and Dessai, 2010).

Uncertainties can be characterized via an ensemble approach, meaning that multiple simulations are run rather than only a single simulation. Following an ensemble approach, one aspect of the modeling chain is adjusted prior to the simulation start. The simulation is then ran and this process is repeated in order to test the range of possible outcomes in response to the adjustment made. From the different simulations, typically the mean or median is calculated as well as the range across all ensemble members, referred to as ensemble spread. This spread is commonly used to interpret the uncertainty of the simulations. Ensembles can be used to test for structural uncertainty by using different model structures and characterizing the range of their output. The structural uncertainty related to the use of different hydrological models is addressed in Paper IV and structural uncertainty related to different GCM-RCMs is addressed in Papers III and IV. Additionally, ensembles can be used to incorporate parametric uncertainty, which in a hydrological modeling context, involves the running of the hydrological model multiple times using different parameter sets (see Paper II, IV, and V). More novel aspects of uncertainty testing were conducted in Paper III, where we analyzed differences in hydrological projections when a univariate versus a multivariate bias correction approach was used. In Paper IV we analyzed the impact of different levels of $E_{pot}$ information on hydrological projections. Although uncertainties have created formidable roadblocks to progress in hydrological climate change impact modeling, uncertainties should not prevent decision making nor the creation of climate change adaptation strategies (Knutti and Sedláček, 2013).

## 1.10 Dissemination of results

There are two main perspectives for carrying out a climate change impact assessment, which have distinctly different modes of disseminating results. By far the most common approach is a 'top-down' approach. This approach is referred to as 'top-down' because the modeling chain is initialized 'from the top', where scientific interests are kept as the main focus and applicability of the study to end-users is only generally addressed. This approach is widely found within the literature, and there are few instances where adaptation strategies or planned decision making has resulted from this approach. In contrast, a 'bottom-up' approach begins with identifying local or end-user vulnerabilities to climate

change impacts, based on a host of factors. Vulnerability indicators (thresholds) are prescribed, which can be helpful in tracking how the risk to climate change exposure develops over time. Through iterative rounds of feedback with end-users, the results of the impact modeling chain can be tailored to support the development of climate change adaptation strategies and support decision making. Given the prospect of ongoing and impending climate change, the argument for committing more resources to adaptation studies that incorporate end-user concerns and vulnerabilities is becoming more relevant (Parry et al., 2009; Wilby and Dessai, 2010; UNDP, 2007).

Paper V of this dissertation delves into this topic more deeply, where we employed a bottom-up approach to carry out a climate change impact assessment. We collaborated with a hydroelectricity company in order to support their adaptation to climate change. Within Switzerland, ownership of public waters is assigned to the cantonal or local/ municipal authorities, who can then grant the right to use the water for specific purposes, known as a concession (Mauch and Reynard, 2004). Concession negotiations play an important role within climate change adaptation given that most dams in Switzerland were built between 1945 and 1970, and water concessions were typically granted for up to 80 years. Therefore, many hydropower managers throughout Switzerland are in the process of renegotiating their water concessions, transforming their existing infrastructure according to updated environmental standards, and considering investments in new locations and alternative energy sectors (Barry et al., 2015). However, unfortunately within the vast majority of concession negotiations, tailored climate change impact analyses are not used (Tonka, 2015). Through our close collaboration with hydropower managers, we were able to draw conclusions regarding the use of a bottom-up approach in climate change impact analyses.

# Chapter 2

# Thesis objectives

## 2.1 Overview

This dissertation delves into the numerical modeling approach to achieve projections of climate change impacts on water resources in order to support adaptation. The five scientific studies presented here sought to understand and summarize the individual components of a typical impact modeling chain (Paper I), rethink evaluation methods of climate models and their bias correction (Paper II), test the sensitivity of hydrological projections when we adjust parts of the modeling chain (Papers III and IV), and apply a climate change impact modeling chain to a real-world decision making problem (Paper V).

## 2.2 Paper I

Hakala, K., N. Addor, C. Teutschbein, M. Vis, H. Dahklaoui, and J. Seibert (2019), Hydrological modeling of climate change impacts. Encyclopedia of Water: Science, Technology, and Society, 1-20, doi:10.1002/9781119300762.wsts0062.

Hydrological climate change impact modeling entails the use of a modeling chain and stepwise procedures, which are not easily applied without extensive instruction and background knowledge. Often an individual modeler is trained over months to years on how to navigate through the different steps of the modeling chain. Given this lengthy orientation period, we organized materials and wrote an encyclopedia chapter to facilitate the training of new or even experienced modelers. The goals of this paper were to review and summarize the steps to perform a climate change impact study, discuss the associated sources of uncertainty, and provide guidance on best practices. An all inclusive summary of these topics was previously missing from the literature. We were especially motivated by the need to facilitate the production of new climate change impact studies and to promote young scientists who endeavor to conduct their own climate change impact assessment.

## 2.3 Paper II

Hakala, K., N. Addor, and J. Seibert (2018), Hydrological modeling to evaluate climate model simulations and their bias correction. Journal of Hydrometeorology, 19, 1321-1337, doi:10.1175/JHM-D-17-0189.1.

In this paper we take another look at the standard format of a climate change impact analysis. When following the typical steps of a modeling chain, variables simulated by GCM-RCMs are evaluated independently prior to using the data as input to a hydrological model. However, the impacts from climate change are often the result of the interactions between variables; these interactions are not considered in current evaluation methods. In this paper, we propose an evaluation method based on the concept that streamflow is controlled by a wide range of hydrometeorological processes, and thus when streamflow is simulated, the realism of the streamflow simulations reflects how well those processes are represented in the models. We therefore present a more process-based evaluation method which simultaneously considers the realism of the individual climate variables, as well as their interaction with one another. This research shows how hydrological modeling represents a novel way to assess the realism and support the selection of climate models.

## 2.4 Papers III & IV

Meyer, J., I. Kohn, K. Stah, K. Hakala, J. Seibert, and A. J. Cannon (2019), Effects of univariate and multivariate bias correction on hydrological impact projections in alpine catchments. Hydrology and Earth System Science, 23, 1339-1354, doi:10.5194/hess-23-1339-2019.

Dahklaoui, H., J. Seibert, and K. Hakala (2020), Sensitivity of discharge projections to potential evapotranspiration estimation in Northern Tunisia. Regional Environmental Change, 20, 1-12, doi:10.1007/s10113-020-01615-8.

Papers III and IV represent efforts to evaluate the sensitivity of the modeling chain to adjustments in its design. We tested system sensitivity in response to changes in the method used for bias correction (used to adjust the representation of intervariable relationships) and the method used to estimate $E_{pot}$. When performing a climate change impact analysis, climate forcing data are usually downscaled and bias corrected by univariate approaches, which neglect the relationships that exist between climate variables. Within Paper III, we test the hypothesis that the explicit consideration of the relationship between climate variables will affect projections of streamflow in a snow-dominated mountain environment. Glacio-hydrological simulations were performed for two partially-glacierized catchments, located within the Swiss Alps. Our results showed that when a multivariate and univariate bias correction method are compared, the choice of bias correction

method translated into considerable consequences for the hydrological responses of the catchments. Differences in simulated total streamflow were found to be negligible, however systematic differences were found in the seasonally delayed streamflow components from snowmelt.

Within Paper IV, we evaluated system sensitivity to adjustments in the level of $E_{pot}$ information provided to a hydrological model. Specifically we simulated the effect of having more versus less $E_{pot}$ information by first estimating $E_{pot}$ with a temperature-based equation and then aggregating the data, thereby reducing the level of $E_{pot}$ information. In this way, our modeling framework was meant to simulate the use of a data-intensive versus a simple $E_{pot}$ equation. Our results demonstrated that projections of streamflow in a semi-arid region are not sensitive to the degree of $E_{pot}$ information fed into the hydrological model. This finding is useful for those who wish to perform a climate change impact analysis in an arid region, where data-intensive $E_{pot}$ estimates add little value. Rather a simple, temperature-based estimate would suffice. This promotes additional climate change impact studies in arid regions with data scarcity issues, given that simple $E_{pot}$ estimates were shown to be valid.

## 2.5 Paper V

Hakala, K., N. Addor, T. Gobbe, J. Ruffieux, J. Seibert (2020), Risks and opportunities for a Swiss hydroelectricity company in a changing climate. Hydrology and Earth System Sciences, 24, 3815-3833, doi: 10.5194/hess-24-3815-2020.

For this dissertation, we endeavored to apply a climate change impact modeling chain to a real-world decision making problem. We utilized a bottom-up approach in order to create tailored hydrological and climatological projections that were designed to support a hydropower company's adaptation to climate change and to support the renegotiation of their water concession. We collaborated with Groupe E, a Swiss hydroelectricity company, and were specifically tasked to analyze the future inflow entering two of Groupe E's reservoirs: (i) the Vernex (Rossiniére) dam and (ii) the Montsalvens dam. Groupe E requested hydrological projections in order to help them gauge the flexibility of their future operations and to discern whether the inflow entering their reservoirs warrants the renewal of their concessions. During concession negotiations, Groupe E representatives stated that the following will be considered: (i) the development of the energy markets and competitors, (ii) the projected supply of water resources, (iii) changes in energy demand, and (iv) costs associated with adhering to new environmental standards. This study focused on the future estimation of Groupe E's water resources (point ii) and also provided preliminary insights into future energy demand (point iii). By following a bottom-up approach we were able to create tailored hydrological and climatological projections that were designed to support a hydropower company's adaptation to climate change.

# Chapter 3

# Data and methods

## 3.1 Study catchments and discharge data

The catchments investigated differ between the dissertation papers. The catchments selected for this thesis have limited human influence and are located upstream from any major hydraulic installation, such as dams or water transfers. Table 3.1 provides a summary of each paper's catchment attributes. The data providers and motivation for the selection of the catchments are described below:

Paper II: catchments were selected to represent a wide range of regime types and elevations. The research catchments for Papers II and III are designated and managed by the Swiss Federal Office for the Environment (FOEN). Daily discharge data (24-h mean) were provided by the FOEN.

Paper III: catchments were selected so that they are located at high-elevation and are both partially glacierized. Both catchments are located in the Swiss Alps, in the headwater of the Rhine River.

Paper IV: catchments are located within Northern Tunisia, and are all situated at low elevation. Observed discharge data was taken from five streamflow gauges, operated by the General Directorate of Water Resources. The catchments are located within an area, which serves a strategic role as a water supplier for the rest of the country.

Paper V: catchments were delineated to represent the drainage area for two reservoirs, used for hydropower production. The daily reservoir inflow was estimated for the period of 2008-2018 by solving the water balance based on variations of the reservoir level, the volume of turbinated water for hydropower production and estimated losses due to evaporation from the reservoir (reservoir losses to groundwater were neglected). Inflow data for the two reservoirs were provided by the hydropower company, Groupe E.

TABLE 3.1: Catchments for each study are provided along with the following attribute information: catchment area, mean elevation, glacier coverage, and karst areas. All papers focused on Swiss catchments, with the exception of Paper IV, which was based on Tunisian catchments. Please note, Paper I did not explicitly work with catchments.

| Paper | Country | Catchment | Area ($km^2$) | Mean elevation (m.a.s.l.) | Glacier coverage (%) | Karst areas (%) |
|---|---|---|---|---|---|---|
| II | Switzerland | Murg-Wängi | 78.9 | 650 | 0 | 0 |
| | | Mentue-Yvonand | 105 | 679 | 0 | 0 |
| | | Guerbe-Belp | 47.4 | 927 | 0 | 0 |
| | | Breggia-Chiasso | 47.4 | 927 | 0 | 95 |
| | | Cassarate-Pregassona | 73.9 | 990 | 0 | 0 |
| | | Sitter-Appenzell | 74.2 | 1252 | 0.08 | 0 |
| | | Allenbach-Adelboden | 28.8 | 1856 | 0 | 8 |
| | | Dischmabach-Davos | 43.3 | 2372 | 2.1 | 0 |
| III | Switzerland | Schwarze Lütschine | 179.9 | 2059 | 16.5 | 0 |
| | | Hinterrhein | 53.9 | 2357 | 7.1 | 0 |
| IV | Tunisia | Melah-Ouchtata | 315 | 423 | 0 | 0 |
| | | Maaden-Bou Brima | 145 | 337 | 0 | 0 |
| | | Joumine-J. Antra | 234 | 375 | 0 | 0 |
| | | El Abid-Ponte-route | 81 | 202 | 0 | 0 |
| | | Rhezala-Fernana | 138 | 470 | 0 | 0 |
| V | Switzerland | Vernex-Montbovon | 398.5 | 1639 | <1 | 15 |
| | | Montsalvens-Broc | 172.7 | 1386 | 0 | 0 |

## 3.2 Observed climate data

Observed climate data were retrieved from different sources depending on the study. Based on data availability, all studies followed the approach shown in Oudin et al. (2005) for $E_{pot}$ estimation. This formula is based on estimated clear daily sky solar radiation and mean daily air temperature. The data and its providers are summarized below:

Papers II and V: observed climate data were retrieved from the 2 × 2 km MeteoSwiss gridded datasets of TabsD (Frei, 2013) and RhiresD (Frei and Schär, 1998; Schwarb, 2000). These gridded products are based on daily temperature (mean of 10-min interval measurements) and precipitation totals measured (automatic and manual) at the high-resolution gauging network of MeteoSwiss, known as SwissMetNet. The effective resolution of RhiresD is roughly 15-20 km or larger, which is the approximate average interstation distance (MeteoSwiss, 2013). Area-weighted mean values of precipitation and air temperature were extracted for the study catchments.

Paper III: observed climate data were derived from an observation-based interpolation product, i.e., the 1 km × 1 km gridded daily air temperature and precipitation datasets from the HYRAS product (Rauthe et al., 2013; Frick et al., 2014). Area-weighted mean values of precipitation and air temperature were extracted for the study catchments.

Paper IV: observed precipitation data were derived from a total of 123 rain gauges, situated in the study areas. The rain gauges are operated by the General Directorate of Water Resources. Eight meteorological stations provided daily temperature data, managed by the National Institute of Meteorology of Tunisia. Climate forcing was generated by spatially interpolating the station data on a 2 × 2 km grid using an inverse distance weighting technique.

## 3.3 Simulated climate data

For Papers II-V, daily temperature and precipitation series were obtained from the Coordinated Regional Downscaling Experiment (CORDEX; www.cordex.org). CORDEX is part of a collaborative modeling effort where GCM projections from the Coupled Model Intercomparison Project (CMIP5; https://cmip.llnl.gov/cmip5/) were downscaled using RCMs operated by different research institutes. Given the locations of the catchments (i.e. Switzerland, Tunisia), GCM-RCMs were selected from the European domain of the CORDEX project (EURO-CORDEX; http://www.euro-cordex.net/). EURO-CORDEX provides simulations at 0.11° (∽12.5km) and 0.44°(∽50km) on a rotated grid. Given the size of the catchments investigated, the higher resolution 0.11° simulations were used for all studies. For additional information regarding EURO-CORDEX, we refer to Kotlarski et al. (2014), which provides an evaluation of ERA-Interim-driven EURO-CORDEX scenarios for Europe.

## 3.4 Bias correction

The bias correction methods employed for Papers II-V follow a similar methodology to one another, but in some cases the methods differ due to specific post-processing needs of the data. In all cases, bias correction is applied to both temperature (T) and precipitation (P) GCM-RCM data. The methods employed did not explicitly differentiate between the biases of the GCMs and RCMs. Rather, the aggregated total bias (RCM bias and remnant bias from the GCM) was corrected. The individual methods for each paper are described below:

Paper II-V: a univariate quantile mapping method (QM) was used, where the differences between the cumulative distribution functions (CDFs) of observed and GCM-RCM T and P are first calculated. Based on these differences, a transfer function was then calculated which in turn was applied to the GCM-RCM data in order to match its quantiles with those of the observed T and P. The daily time series of T and P were broken up into seasons: December-February (DJF), March-May (MAM), June-August (JJA), and September-November (SON) for both the observed and the GCM-RCM-simulated climate variables. The "qmap" package in R (Gudmundsson et al., 2012; Gudmundsson, 2016) was used to map the CDF of the simulations onto the CDF of the observations.

Paper III: two different bias correction methods were applied to each climate model's T and P series: a univariate quantile mapping technique (QDM; referred to above as 'QM' but this paper utilizes a different acronym for the same term used in the other dissertation papers) and a multivariate bias correction approach (MBCn). The MBCn algorithm by Cannon (2018) is based on the N-dimensional probability density function transform. This MBCn approach combines univariate quantile mapping with random orthogonal rotations to match the multivariate distributions of climate model data and observed data. In the MBCn approach, a random orthogonal rotation of the data points is applied before applying univariate quantile mapping. This exposes the univariate-corrected data to a linear combination of the original variables, which is then used to correct the marginal distributions of the rotated data. The corrected dataset is then rotated back and convergence to the observed multivariate distribution is checked.

## 3.5 Hydrological modeling

For Papers II-V, the bucket-type Hydrologiska Byråns Vattenbalansavdelning (HBV) model (Bergström, 1976; Lindström et al., 1997) was used to simulate daily streamflow values (please note that no hydrological simulations were conducted for Paper I). Specifically, the version HBV-light (Seibert and Vis, 2012) was used, referred hereafter to as HBV. The basic model structure of HBV is shown in Figure 3.1. HBV was developed at the Swedish Meteorological and Hydrological Institute (SMHI) in the 1970s. HBV requires daily temperature, precipitation, and $E_{pot}$ values as driving variables. The HBV model relies on four routines, that are meant to represent (i) snow pack dynamics, (ii) soil moisture variation, (iii) runoff response, and (iv) discharge routing. For Papers II-V, HBV was applied in a semi-distributed form by disaggregating each catchment into elevation zones based on a digital elevation model. During a preliminary analysis, it was noted that observed lapse rate and precipitation gradient values significantly deviated from HBV's default values (temperature lapse rate of -0.6 C° $(100m)^{-1}$ and precipitation gradient of 10% $(100m)^{-1}$). Therefore, catchment specific long-term mean monthly values were used.

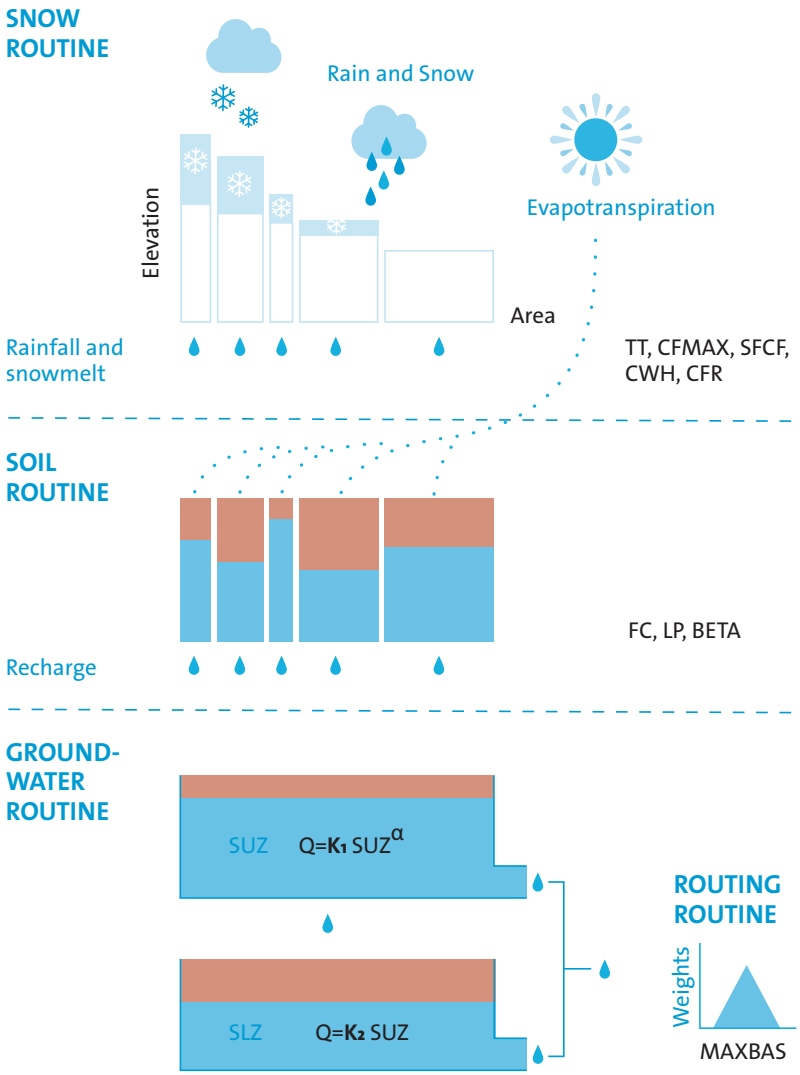



FIGURE 3.1: Structure of the HBV model, adapted from Seibert and Vis (2012) and Bergstrom (1992).

For Paper IV, discharge simulations were created using different model structures. Three simple bucket-type hydrological models were used: Modèle du Génie Rural à 4 paramètres Journaliers (GR4J) (Perrin et al., 2003), HBV (described above), and Identification of unit Hydrograph and Component flows from Rainfall, Evapotranspiration and Streamflow (IHACRES) (Jakeman et al., 1990). These particular models were chosen because they differ in the way that they compute $E_{act}$ from $E_{pot}$, which allows for an inter-comparison of model simulations. Further details of these models can be found within Paper IV.

## 3.6 Extraction of climate variables

For Papers II-V, daily simulated temperature and precipitation time series were extracted from EURO-CORDEX GCM-RCM gridded output. Six different extraction methods were compared in Paper I, all showing similar results. Given its relative simplicity, the area-weighted mean method was chosen, which includes the following steps: (i) all grid cells that overly the catchment shapefile were identified, (ii) individual grid cells were given particular weights based on their percentage overlap with the catchment shapefile, (iii) the precipitation and temperature time series were extracted from the overlying grid cells, and (iv) the average of all extracted time series was calculated, resulting in an area-weighted mean time series. In the same manner, observed climate variables were extracted from their respective gridded data source (see Section 3.2). Both simulated and observed climate variables were extracted for the catchments listed in Table 3.1.

## 3.7 Analysis of simulations

The method used for the analysis of the projections was based upon the individual motivations for each study (described within Chapter 2). A brief recap of each paper's motivation is included along with a description of the associated analysis of the projections.

Paper II: Since streamflow inherently incorporates the dynamics between temperature and precipitation at the catchment scale, the evaluation of simulated discharge, with and without bias correction, can be used to determine if the the relationship between climate variables is properly represented by a given climate model. We used this concept to set up our modeling framework. Over the historical time period, the following streamflow series were created:

- $Q_{obs}$: Observed discharge
- $Q_{ref}$: Simulated discharge (hydrological model forced by observed climate data)
- $Q_{raw}$: Simulated discharge (hydrological model forced by raw GCM-RCM data)
- $Q_{qm}$: Simulated discharge (hydrological model forced by bias-corrected GCM-RCM data)

From this it follows that the differences between the above mentioned streamflow series can be used to represent errors in the modeling chain. Differences between the streamflow series can reflect errors in the following way:

- $Q_{obs}$ vs $Q_{ref}$: errors in the observed climate forcing dataset as well as limitations in the hydrological model
- $Q_{ref}$ vs $Q_{raw}$: errors resulting from GCM-RCM biases
- $Q_{raw}$ vs $Q_{qm}$: the impacts of bias correction

For all streamflow time series, these differences were analyzed according to several discharge metrics (i.e. DJF mean, JJA mean, Q5, Q95, 7-day low flow, annual maximum, half-flow date). These metrics were chosen as they are typical metrics used for water management purposes. To standardize the metrics, the relative error was calculated by comparing $Q_{ref}$ to $Q_{obs}$, $Q_{raw}$, and $Q_{qm}$ in the following way:

- $E_{obs}$= ($Q_{obs}$ - $Q_{ref}$)/$Q_{ref}$
- $E_{raw}$= ($Q_{raw}$ - $Q_{ref}$)/$Q_{ref}$
- $E_{qm}$= ($Q_{qm}$ - $Q_{ref}$)/$Q_{ref}$

According to these standardized metrics, the different climate models were ranked before and after bias correction. The ranking was used to interpret the relative realism of each GCM-RCM's climate variables and their intervariable relationships.

Paper III: The choice of bias correction method within a climate change impact modeling chain can influence important dynamics related to hydrological projections. In this way, bias correction represents a source of uncertainty in the modeling chain. High elevation catchments are especially sensitive to changes in climate as they lie within the elevation range where snow accumulation occurs. Given this delicate relationship between temperature and precipitation, it is especially critical to simulate the interdependence between these two variables correctly. Thus, two bias correction approaches, MBCn and QDM (described in detail within Section 3.4), were compared to one another. The simulation results were assessed in terms of: snow water equivalence (SWE), glacier ice volume (V) evolution, and streamflow with three individual components:

- $Q_R$: streamflow resulting from rainfall
- $Q_S$: streamflow resulting from snow melt
- $Q_I$: streamflow resulting from ice melt

Once the bias correction was validated over the historical period (1976-2006), the two methods were compared according to their influence on the projections (2006-2099) of streamflow, SWE, and V.

Paper IV: The estimation of $E_{pot}$ is an another source of uncertainty within a hydrological climate change impact analysis. On one hand, the use of physically-based $E_{pot}$ equations may result in more realistic hydrological projections compared to simpler temperature-based equations, however the required data is often not available to make this full comparison. Therefore, this study employed a novel approach to test the sensitivity of hydrological projections to $E_{pot}$ estimation, within a data limited region. In this way, we were able to gauge the impact of using a simple vs physically-based $E_{pot}$ equation within a typical climate change impact study. To test this we utilized the Oudin et al. (2005) formula for estimating $E_{pot}$, where $E_{pot}$ is estimated in the following way:

$$E_{\text{pot}} = \frac{R_{\text{e}}}{\lambda\rho}\frac{T+5}{100} \text{ if T > -5, otherwise } E_{\text{pot}} = 0$$

Where:
$E_{pot}$: the rate of potential evapotranspiration (mm $day^{-1}$)
$R_e$: extraterrestrial radiation (MJ $m^{-2}$ $day^{-1}$)
$\lambda$: latent heat flux (MJ $kg^{-1}$)
$\rho$: density of water (kg $m^{-3}$)
T: mean air temperature (°C)

Using the Oudin formula, we created daily mean values of $E_{pot}$. Given the lack of data required to fulfill more physically based equations, we instead reduced our level of $E_{pot}$ information to infer the influence of varying levels of $E_{pot}$ information. Following this concept, $E_{pot}$ was calculated in three different ways:

- Case 1: Daily $E_{pot}$ calculated over the calibration and validation period separately (without averaging).
- Case 2: Long-term daily mean $E_{pot}$ with the same values for the calibration and validation periods ($E_{pot}$ calculated over the calibration period and then averaged to create a series of 366 values, which were then repeated for both the calibration and validation period).
- Case 3: Long-term daily mean $E_{pot}$ varied between calibration and validation periods ($E_{pot}$ calculated over the calibration period and validation period separately. Each set of daily values were then averaged to create two sets of 366 values, which were then used for calibration and validation separately).

The three cases described above were compared for specific purposes and a diagram of the three cases is shown in Figure 3.2. Over the historical period, the meaning behind each case is rather straightforward. A comparison between Case 1 and 3 allowed for an evaluation of the impact of long-term averaging of $E_{pot}$ on the hydrological model performance. A comparison of these two cases was also performed in order to determine whether long-term averaging of $E_{pot}$ is acceptable for hydrological projections. A comparison between Cases 2 and 3 represented the impact of keeping $E_{pot}$ the same versus allowing it to change between calibration and validation on hydrological model performance. Over the future period, Case 1 is meant to represent a situation where a simple temperature-based $E_{pot}$ formula is used. Case 2 represents a situation, similar to the pan evaporation paradox (see Chapter 7 for further discussion of this paradox), which expects an attenuation of future $E_{pot}$ by the change in other climate drivers, resulting in stabilization. This kind of attenuation can only be addressed by physically-based $E_{pot}$ formula. Case 1 and Case 2, in this instance, are meant to represent two extremes of $E_{pot}$ estimation. Thus, a comparison between these two cases allows for the evaluation of the sensitivity of hydrological projections to $E_{pot}$ estimation.

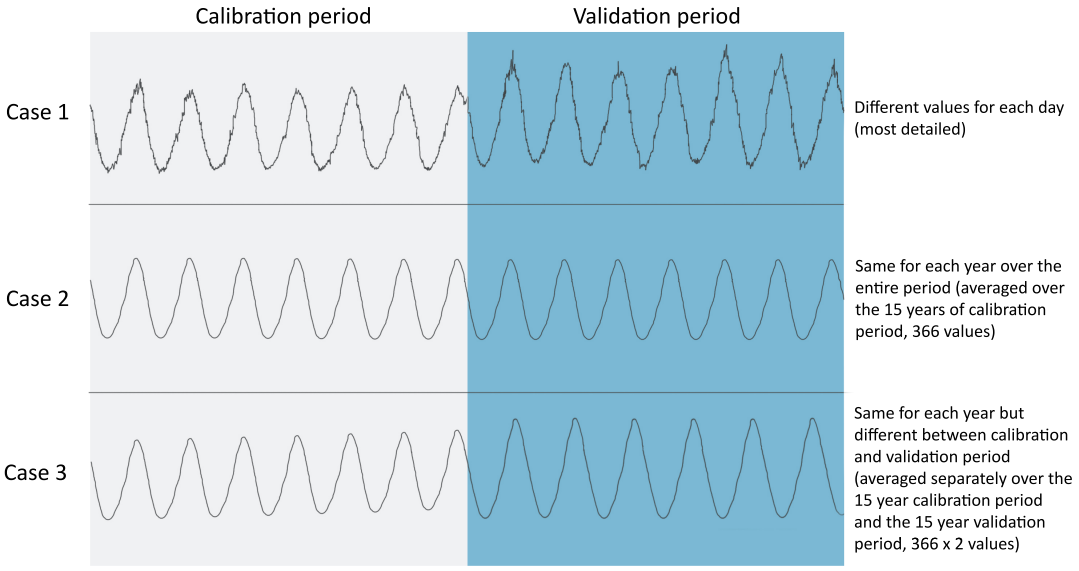



FIGURE 3.2: Diagram of the different $E_{pot}$ formulations used in the study, including: (i) Case 1: different daily values, (ii) Case 2: 366 daily repeating values over both calibration and validation period, calculated by averaging over the 15 year calibration period, and (iii) Case 3: 366 daily values, calculated and repeated over calibration period and 366 different values calculated and repeated for the validation period. Figure from Paper IV.

Paper V: In this study we collaborated with Groupe E, a hydroelectricity company in Switzerland. Over several meetings, we met with Groupe E representatives and asked them to identify which hydroclimatic changes their operations are most vulnerable to. Together we then chose streamflow and energy demand indices that were meant to characterize their associated risk under future climate. Groupe E also specified thresholds in relation to the streamflow indices which, if exceeded, would represent a significant impact to their operations. Table 3.2 summarizes the streamflow indices, thresholds, and describes the relevance of each index to Groupe E's operations.

In addition to the streamflow indices, two indices were used to gauge energy demand, namely cooling degree days ($C_{DD}$) and heating degree days ($H_{DD}$), calculated in the following way:

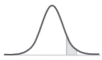
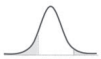
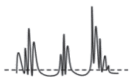
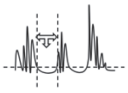
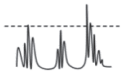
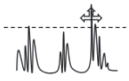
$$H_{DD} = \max(\text{Th} - \theta\text{t}, 0)$$

$$C_{DD} = \max(\theta\text{t} - \text{Tc}, 0)$$

Where:
$\theta$t is the air temperature extracted from the GCM-RCMs. The threshold Th= 13°C and Tc= 18.3°C were provided by Groupe E and represent the air temperatures, which if reached, would induce customers to turn on either heating or cooling in their homes, respectively.

$H_{DD}$ and $C_{DD}$ were calculated for the cities (cantonal boundaries) of Zurich and Geneva, as these areas represent the locations of typical Groupe E customers. $C_{DD}$ and $H_{DD}$ were computed following the method presented in Gaudard et al. (2013).

TABLE 3.2: Hydrological indices selected after discussions with representatives of Groupe E. The relevance of each index for Groupe E's operations is explained, and vulnerability thresholds for each index are provided. Relative changes exceeding these thresholds would have a significant impact on Groupe E's operations. In cases where two thresholds are provided, the exceedance of the lower threshold represents a significant impact and the upper threshold represents a critical impact. Table from Paper V

| Category | Hydrological index (season) | | Specific relevance for Groupe E | Vulnerability thresholds | Visual |
|---|---|---|---|---|---|
| Water volume | Long-term seasonal mean | March, April, May (MAM) | Snowmelt runoff may coincide with high intensity precipitation events that could overwhelm Groupe E's ability to store and turbinate reservoir inflows. | 20%, 50% |  |
| | | June, July, August (JJA) | Meeting water demand for recreation, esthetic, and residual flow requirements. | |  |
| | | September, October, November (SON) | Managing reservoir level given drought concerns, meeting demand for recreation, esthetic, and residual flow requirements. | 20%, 50% |  |
| | | December, January, February (DJF) | Meeting energy demand during the coldest time of year, and identifying opportunities to benefit from historically high energy prices. | |  |
| Low flows | Q5: 5th percentile of daily streamflow | June, July, August (JJA) | Meeting water demand for recreation, esthetic, and ecological purposes. Important for water fee negotiations and to assess whether regulations for residual flows are realistic. | 50% |  |
| | | September, October, November (SON) | Meeting water demand for recreation, esthetic, and ecological purposes. Important for water fee negotiations and to assess whether regulations for residual flows are realistic. | | |
| | Consecutive days of low flow | June, July, August (JJA) | Help with storage management during extended drought, concern for meeting energy demand and residual flow requirements. | 60 days |  |
| | | September, October, November (SON) | Storage management during extended drought, concern for meeting energy demand and residual flow requirements. | | |
| High flows | Q95: 95th percentile of daily streamflow | June, July, August (JJA) | Reservoir levels are at their highest, multi-day high intensity precipitation can lead to water release without turbination (profit loss) or damage downstream in extreme cases. | 50% |  |
| | | December, January, February (DJF) | Explore opportunities to benefit from high energy prices during the winter months, although DJF has not produced as much inflow as JJA. | | |
| | Consecutive days of high flow | June, July, August (JJA) | Reservoir level management and balancing dam releases with high intensity precipitation events. | 10 days |  |
| | | December, January, February (DJF) | Utilizing high volumes of inflow at times when the unit price of water is its highest. | | |

## 3.8 Design of the modeling chain

The main structure for the experimental design, as described in the introduction, was similar for Papers II-V. However, the individual attributes of each modeling chain differed in its selection of (i) study catchments, (ii) time periods and emission scenarios, (iii) climate models, (iV) post-processing methods, and (V) hydrological model setup and parameterization. The attributes for each modeling chain are presented in Table 3.3.

TABLE 3.3: Catchments for each study are provided along with the following attribute information: catchment area, mean elevation, glacier coverage, and karst areas. Please note, Paper I did not utilize a modeling chain.

| | Paper II | | Paper III | | Paper IV | | Paper V | |
|---|---|---|---|---|---|---|---|---|
| Study sites | 8 Swiss catchments | (×8) | 2 Swiss catchments | (×2) | 5 Tunisian catchments | (×5) | 2 Swiss catchments | (×2) |
| Time frame | 1970-2009(-) | (-) | 1976-2006(-)<br>2006-2099 | (×1) | 1970-2000(-)<br>2040-2070<br>2070-2099 | (×2) | 1980-2009(-)<br>2006-2099 | (×1) |
| Emission scenarios | Historical only | (-) | RCP4.5<br>RCP8.5 | (×2) | RCP4.5<br>RCP8.5 | (×2) | RCP4.5<br>RCP8.5 | (×2) |
| Climate models | 12 GCM-RCMs | (×12) | 10 GCM-RCMs | (×10) | 11 GCM-RCMs | (×11) | 11 GCM-RCMs | (×11) |
| Post-processing | Raw data<br>1 bias correction method | (×2) | Raw data<br>2 bias correction methods | (×3) | 1 bias correction method | (×1) | Raw data<br>1 bias correction method | (×2) |
| Hydrological model | HBV | (×1) | HBV | (×1) | HBV<br>GR4J<br>IHACRES | (×3) | HBV | (×1) |
| Objective function | $R_{Lindström}$ | (×1) | Weighted objective function:<br>$R_{eff}$<br>$R_{KGE}$<br>$R_{VE}$<br>$R_{Lindström}$<br>$R_{eff\_log}$<br>$R_{eff\_JS}$<br>RMSE (Snow)<br>$MAE_{Normalized}$ (SWE)<br>Glacier volume change | (×1) | Differential split sampling was performed for testing using the following:<br><br>(Calibration)<br>$R_{KGE}$<br><br>(Validation)<br>$R_{eff}$<br>$R_{VE}$ | (×1) | $R_{eff}$<br>$R_{Lindström}$<br>$R_{KGE}$ | (×3) |
| Time periods calibration\ (validation) | 1980-1994<br>(1995-2009) | (×1) | 1976-2003<br>(2003-2006) | (×1) | Full time period used to generate one parameter set validation done only for testing:<br><br>1970-2000 | (×1) | 2008-2013<br>(2013-2018)<br>2013-2018<br>(2008-2013) | (×2) |
| Calibration method (# parameter sets) | GAP<br>(10 sets) | (×10) | GAP<br>(1 set) | (×1) | Shuffle Complex Algorithm<br>(1 set) +<br>2 $E_{pot}$formulations | (×2) | GAP<br>(10 sets) | (×10) |
| No. of GCM-RCM forced streamflow projections | 1920 | | 120 | | 1320 | | 7920 | |

*For abbreviation definitions, please see Acronyms (page vii) in the front matter.*

# Chapter 4

# Results

## 4.1 Evaluating climate models and their bias correction

### 4.1.1 Performance of bias correction

The first objective of Paper II was to test how meteorological intervariable relationships are affected by bias correction and how do these dynamics influence the hydrolgical simulations. Prior to bias correction, biases in raw GCM-RCM precipitation ($P_{raw}$) were prominent. The biases generally took the form of either a wet bias that persists throughout the year (Figure 4.1 a) or a wet bias that generally occurs in the winter and spring months, with an associated dry bias in the summer months (Figure 4.1 b,c). These biases are representative of similar biases found in other catchments not shown here. As expected, bias corrected precipitation ($P_{qm}$) showed a strong improvement compared to ($P_{raw}$), according to their relative fit with observed precipitation ($P_{obs}$). However, in some cases, bias can still remain or even degrade after bias correction. Temperature biases were also present, however these biases appeared to be especially sensitive to the elevation. Prior to bias correction, the most apparent biases in temperature ($T_{raw}$) could be seen in high-elevation catchments. In such catchments, a cold bias was found to persist throughout the entire annual cycle. QM improved these cold biases, bringing ($P_{qm}$) and ($P_{obs}$) closer together, irrespective of the elevation of the catchment. The biases in temperature and precipitation had a direct impact on hydrological simulations, however the dynamics were strongly influenced by the elevation of the catchment. Within raw GCM-RCM forced simulations, high elevation catchments (1856-2372m; e.g. Allenbach, Dischmabach) experienced a combination of wet biases in winter/spring as well as a pervasive cold bias, which led to a delay in discharge of 1-1.5 months compared to $P_{obs}$. The delay in discharge is the result of precipitation falling as rain rather than snow, which has a strong impact on the dynamics of high elevation catchments. Bias correction was shown to correct the cold biases at all elevations, thus, snow accumulation and timing of discharge were generally improved after QM. In medium to low elevation catchments (679-1252m; e.g. Sitter, Murg), wet biases were found to either exist only in the winter months or over the entire annual cycle. This translated into a bias in the magnitude of discharge, however timing was not effected.

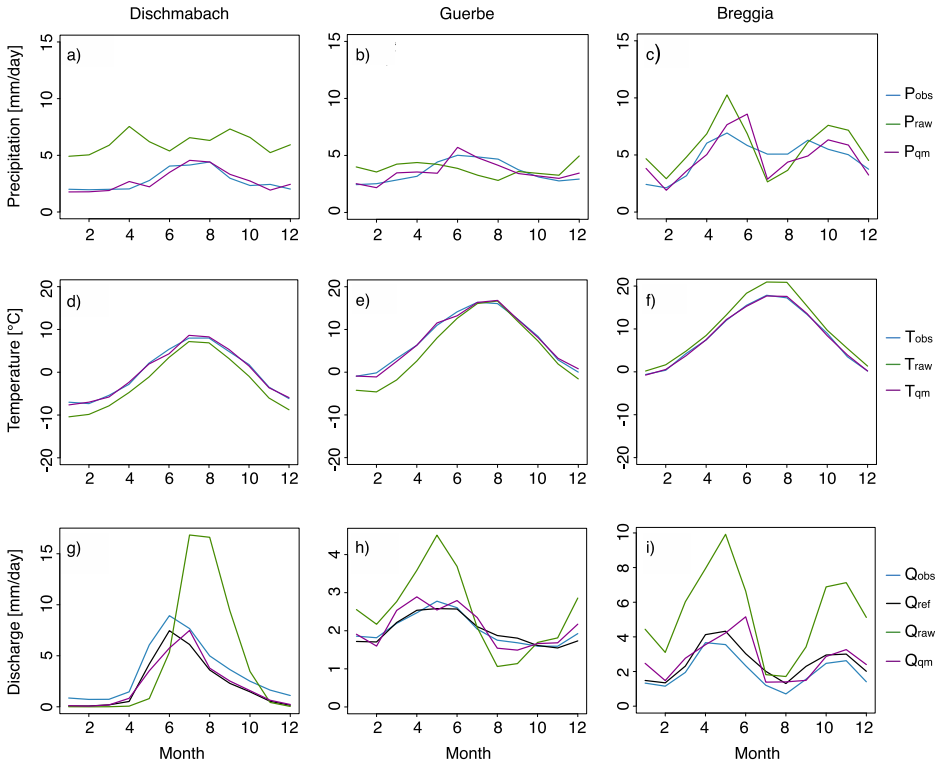



FIGURE 4.1: The long-term mean monthly (a)-(c) precipitation, (d)-(f) temperature, and (g)-(i) discharge for three example catchments. The data from one GCM-RCM are used for each catchment. (left) Dischmabach catchment, CNRM-CM5-CLM4-8-17; (center) Guerbe catchment, CNRM-CM5-RCA4; and (right) Breggia catchment, HadGEM2-ES-RCA4. All figures are for the period 1980-2009. Note the different axis for the three discharge plots in (g)-(i). Figure from Paper II.

A ranking of GCM-RCMs based on $E_{raw}$ and $E_{qm}$ was created according to each metric (Figure 4.2), where $E_{raw}$ and $E_{qm}$ represent the median of all catchments. In Figure 4.2 a, observed discharge can be seen to rank high compared to all raw GCM-RCMs. In contrast, Figure 4.2 b demonstrates that after bias correction, the GCM-RCM data generally fits more closely to $QE_{ref}$, resulting in observed discharge ranking last. Raw GCM-RCMs were found to have more variability amongst themselves, with errors ranging between 26% and 88%. In contrast, QM GCM-RCM forced simulations range between 4% and 11%.

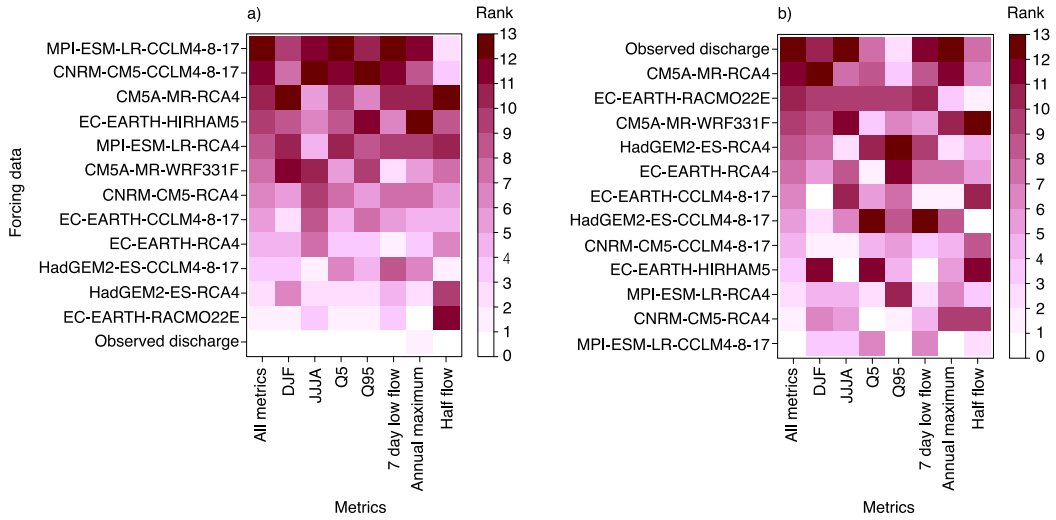



FIGURE 4.2: (a) Raw and (b) QM GCM-RCM (y axis) ranked according to their performance for various hydrological metrics (x axis) across all catchments. The placement of the GCM-RCMs along the y axis is determined by their rank within the All metrics column. Note that observed discharge ranks high in (a) and ranks low in (b). Figure from Paper II

## 4.2 Sensitivity of the modeling chain to bias correction

### 4.2.1 Climate variables and bias correction

Two different bias correction methods (i.e. QDM and MBCn) were applied to a hydrological climate change impact modeling chain within Paper III. The choice of bias correction method led to differences concerning the intervariable relationship between P and T. This is most apparent when comparing the distribution of annual precipitation sums when air temperatures are either above or below 0°C (Figure 4.3). Precipitation sums above and below 0°C can be directly linked to snow accumulation and melt processes. As Figure 4.3 shows, uncorrected climate model data (noBC) is highly variable compared to the reference data (HOCD). In relation to HOCD, precipitation falling above air temperatures of 0°C was overestimated by QDM. Correspondingly, precipitation falling below air temperatures of 0°C was underestimated by QDM. In general MBCn appears to have better reproduced the reference data over the historical period.

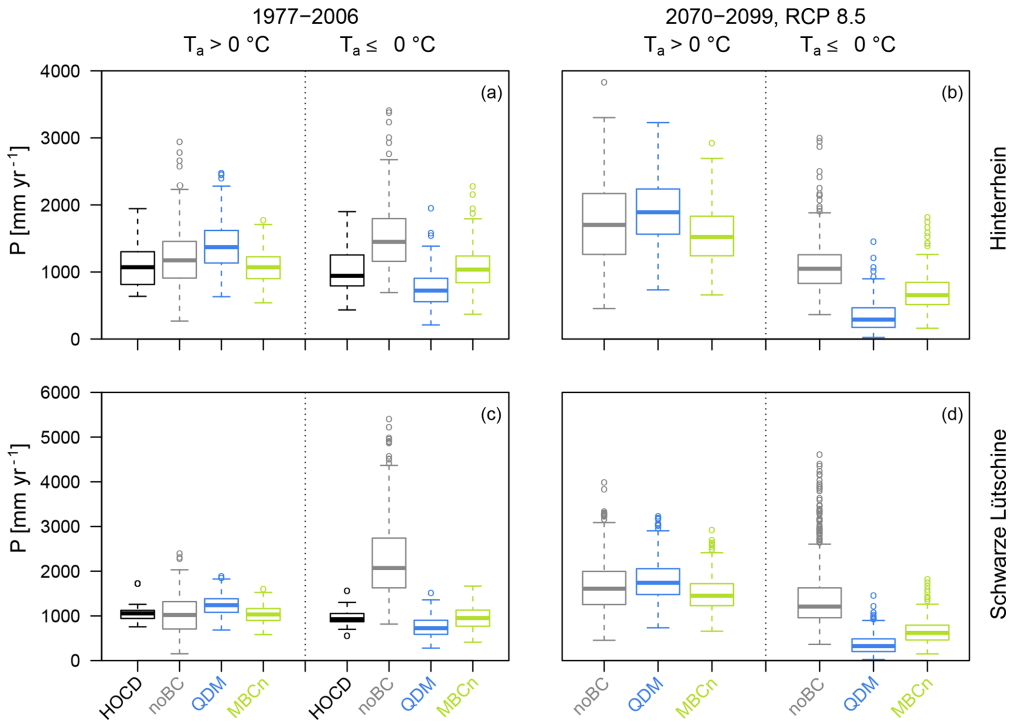



FIGURE 4.3: Annual precipitation sums for days with air temperature above or below 0°C. Figure from Paper III

### 4.2.2 Hydrological simulations

The relationship between temperature and precipitation after the application of QDM versus MBCn to the GCM-RCM data translated into different accumulation amounts of snow water equivalence (SWE). Over the historical period, MBCn corrected GCM-RCM simulations resulted in more precipitation falling as snow, with a greater accumulation of SWE in spring (100-200mm higher). When looking at future simulations, irrespective of the bias correction method applied, snow melt was shown to play a progressively smaller role over the coming century (Figure 4.4). Glaciers in both catchments were also found to continuously diminish over the coming century. To illustrate these points more clearly, the different components of streamflow were separated out as follows:

- $Q_R$: streamflow resulting from rainfall
- $Q_S$: streamflow resulting from snow melt
- $Q_I$: streamflow resulting from glacier ice melt

Figure 4.4 shows that that total annual streamflow will stay generally unchanged over the coming century, however the composition of streamflow will clearly change. The streamflow component for glacier ice melt ($Q_I$) is shown to decrease over time for both catchments, due to glacier retreat. Likewise, the streamflow component from snow melt ($Q_S$)

is also shown to decrease over time. Simulations based on QDM-corrected data led to slightly different total streamflow compared to MBCn-corrected data. However, differences between the bias correction approaches were much more apparent regarding the individual streamflow components.

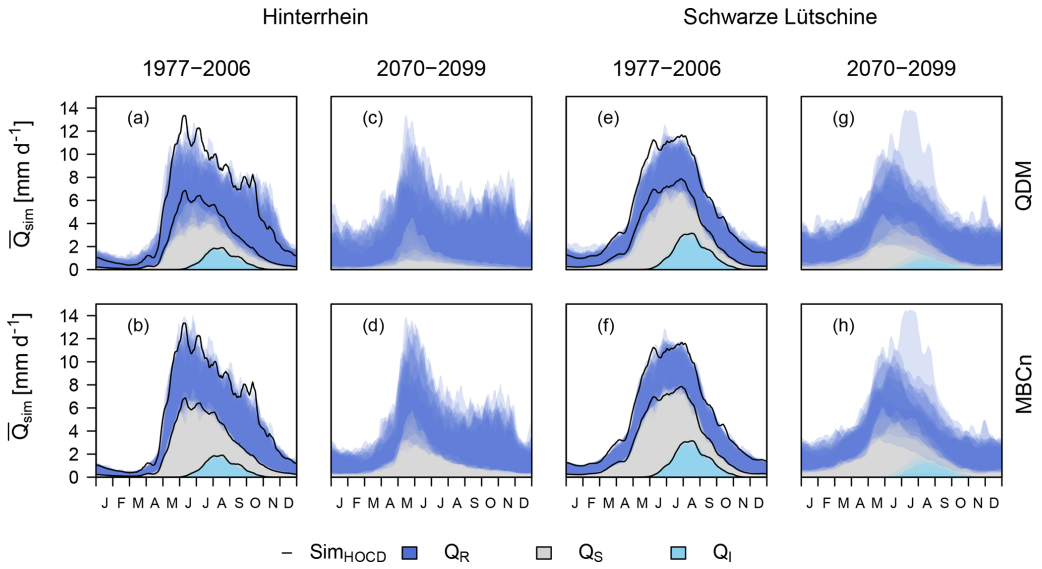



FIGURE 4.4: Streamflow regimes based on 11-day moving averages of daily streamflow during 30-year periods in the historical reference period and as projected for the period 2070-2099 under the RCP8.5 scenario for the two catchments. Simulation results for each ensemble member are shown as semi-transparent polygons. For the historical reference period the reference period the results of the simulations based on the historical reference P and T time series are also shown (black lines). Figure from Paper III.

## 4.3 Sensitivity of the modeling chain to $E_{pot}$

### 4.3.1 Effect of long-term averaging of $E_{pot}$ information on hydrological model performance

The results of Paper IV show that when daily calculated $E_{pot}$ (Case 1) is compared to long-term averaged daily $E_{pot}$ (Case 3), only minimal differences exist. The most significant differences occur during the summer period, when streamflow is at its lowest. Given that these two cases exhibit similar patterns, this implies hydrological simulations, when forced with temporally varying $E_{pot}$ versus long-term daily mean $E_{pot}$, will result in similar streamflow time series. We observed no significant difference between the different hydrological models (i.e. HBV, IHACRES, GR4).

### 4.3.2 Effect of $E_{pot}$ estimation on hydrological model performance

An additional experiment was performed over the historical time period to assess the change in hydrological model performance when $E_{pot}$ is either kept constant between calibration and validation (Case 2) versus allowed to change (Case 3). Our results show that only small differences were found between the different hydrological models. In terms of volume error, the differences ranged between -4% and +4%. For $R_{eff}$, the difference between the two cases was essentially zero.

### 4.3.3 Future simulations

By the end of the century, results of mean annual discharge show that Case 2 is approximately -2 to -8% smaller than Case 1. The differences between the two cases was found to be relatively stable irrespective of which emission scenario was being considered. Similar results were also found amongst the different hydrological models. An analysis of variance (ANOVA) was performed in order to decompose the projection variance among four sources of uncertainty: hydrological models, $E_{pot}$ formulation, emission scenarios, climate models (GCM-RCMs), interactions amongst the sources of uncertainty, and residual errors. The results of the ANOVA showed that GCM-RCMs were the dominant source of uncertainty, whereas $E_{pot}$ formulation contributed a relatively small amount of uncertainty to the hydrological projections (Figure 4.5).

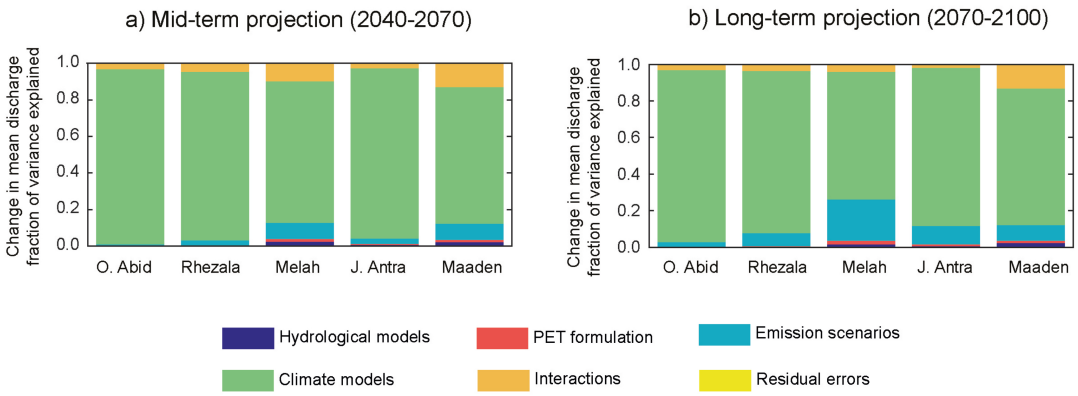



FIGURE 4.5: Decomposition of the projection variance. ANOVA partitioning among the four sources of uncertainty, the significant interactions, and the residual errors for discharge change are shown for the two 30 year future periods (2010-2070 and 2070-2100). Figure from Paper IV.

Hydrological projections also showed that the hydrological response of all catchments was highly sensitive to any change in precipitation. Figure 4.6 shows the elasticity of discharge to precipitation for the hydrological model GR4. Elasticity here is defined as the change in discharge as a function of the change in precipitation. Our results show that a change in precipitation results in a multiplication of discharge by a factor of around two. In other

words, for every unit of precipitation increase or decrease, we can expect that the hydrological response be double the amount. The elasticity curves were found to be similar between Cases 1 and 2 as well as across the different hydrological models.

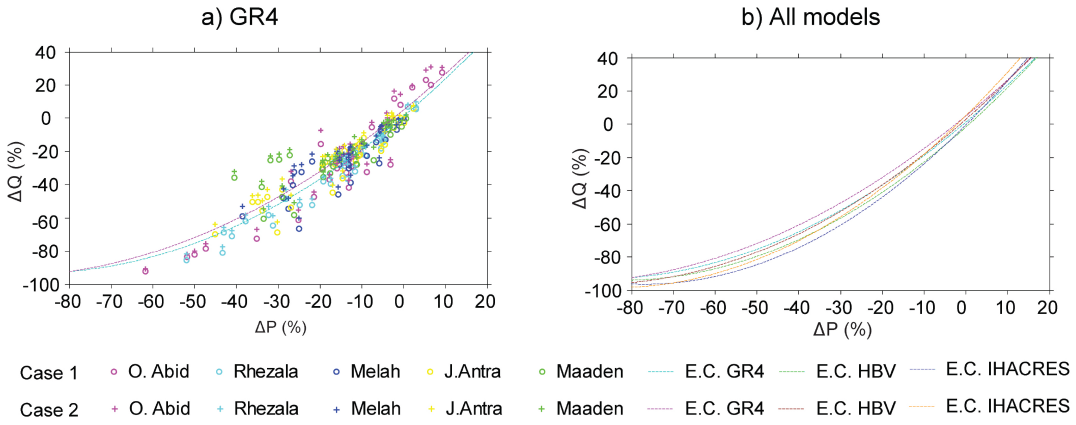



FIGURE 4.6: Elasticity of discharge due to precipitation change for a) GR4 model, and b) show elasticity curves from the other hydrological models. ΔP and ΔQ represent the changes in discharge and precipitation from different rainfall-runoff models, catchments, representative concentration pathway scenarios (RCPs), climate models (GCM-RCMs), time periods, and potential evapotranspiration estimation cases. Figure from Paper IV.

## 4.4 Application of modeling chain to aid decision making

### 4.4.1 Water Volume

Our projections found within Paper V show that winter inflow entering the Vernex reservoir under RCP 8.5 will widely exceed the +20% and 50% thresholds specified by Groupe E. Conversely, inflows over the summer period will decrease substantially, with the distribution over these months residing along Groupe E's -50%, threshold. By the end of the century, the average change of inflow entering the Vernex reservoir will be -1.11 M $m^3$/day (-4.52 to +2.54 likely range) under RCP 8.5 and -0.24 M $m^3$/day (-2.968 to +2.3487 likely range) under RCP 4.5. In the same way, inflow entering the Montsalvens reservoir was also found to undergo a similar redistribution of its inflow, with an average decrease of -0.724 M $m^3$/day (-2.19 to +0.81 likely range) under RCP 8.5 and -0.18 M $m^3$/day (-1.61 to + 1.08 likely range) under RCP 4.5. In all cases, the likely range represents represents two thirds of all 660 simulations per catchment.

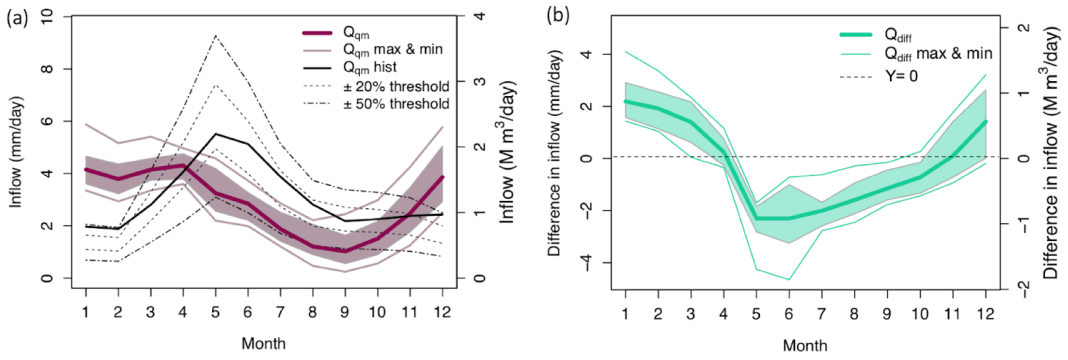



FIGURE 4.7: (a) Long-term mean monthly inflow entering the Vernex reservoir for 1980-2009 ($Q_{qm}$ hist) and for 2070-2099 ($Q_{qm}$, RCP 8.5). The mean (solid lines) and likely range (shaded areas) are shown, where the likely range represents two thirds of all simulations. The two thresholds are based on the mean of the simulations forced by observed climate data ($Q_{ref}$ over the period of 1980-2009). (b) Long-term mean monthly change in inflow (2070-2099 with respect to 1980-2009) for the Vernex catchment. Figure from Paper V.

The distribution changes in streamflow are induced by a shift in the form of precipitation contributing to inflow (Figure 4.8). The annual contribution of snowmelt to streamflow ($S_{qm}$) is expected to decrease by more than half and to occur earlier in the year, shifting from May to April. The contribution of rain to streamflow ($R_{qm}$) will meanwhile decrease over the summer. However, winters will see an increased contribution from rain by the end of the century. The Montsalvens catchment is expected to undergo a similar regime change.

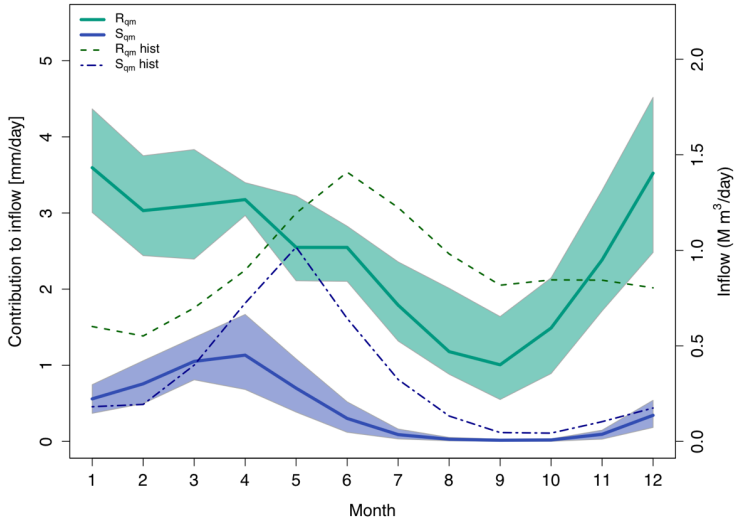



FIGURE 4.8: Mean monthly contribution of rain (R, green) versus snow (S, blue) to inflow entering the Vernex reservoir. Two periods are compared: 1980-2009 ($R_{qm}$ and $S_{qm}$ hist) and 2070-2099 ($R_{qm}$ $S_{qm}$). All projections shown are simulations under RCP8.5. The mean (solid lines) and likely range (shaded areas), are shown, where the likely range represents two thirds of all simulations. The dashed lines indicate the mean of the reference simulations. Figure from Paper V.

### 4.4.2 Low flows

Bias corrected, $Q_{qm}$, simulations of low flows (Q5) for the JJA and SON seasons showed a strong decrease ($\backsim$ -50%) under RCP 8.5. The spread of the ensemble for both seasons is relatively small in absolute terms. Projections for the inflow entering the Montsalvens reservoir indicate similar changes, with the magnitude of Q5 dropping below the -50% threshold for JJA and the median of the SON ensemble lying close to the -50% threshold.

Results also show that the frequency of consecutive days below Q5 will increase by the end of the century over the SON season. Although simulations over the historical period showed an overestimated bias when attempting to replicate observed patterns, yet the robust nature of the change demonstrates that there is high confidence that there will be more days below Q5 over the SON season in the future. The results for the Montsalvens reservoir agree with the changes shown for the Vernex reservoir, with a slightly less pronounced difference between the historical and future periods.

### 4.4.3 High flows

For the Vernex reservoir, the magnitude of of high flows (Q95) was found to decrease over the JJA season under RCP 8.5, but staying within the ± 50% interval, specified by Groupe E. However, for the DJF season, projections of $Q_{qm}$ show a strong increase, surpassing the +50% threshold. Simultaneously, consecutive days above Q95 are expected to increase over the DJF season under the influence of climate change, exceeding the 10 day threshold specified by Groupe E. Montsalvens is expected to experience a similar but less pronounced change.

### 4.4.4 Energy demand

Projections of $H_{DD}$ were found to decrease over the winter months under the influence of climate change, while the summer months will experience no change in $H_{DD}$ given that these months are already warm enough to not induce the need for heating. On the other hand, projections of $C_{DD}$ show an increase for the months between May and October under the influence of climate change. The winter months are not likely to experience any change given that these months are sufficiently cold so that additional cooling is not necessary.

# Chapter 5

# Discussion

## 5.1 Evaluating climate models and their bias correction

Typically within a hydrological climate change impact modeling chain, variables simulated by GCM-RCMs are evaluated independently, where the performance of GCM-RCM temperature data are compared with observation and separately the same is done for precipitation. However this methodology does not consider nor analyze whether intervariable relationships remain intact after bias correction. Within the literature, this limitation is usually overlooked and no further evaluation is performed over the historical period. Typically all available models are carried forward in the chain. This ensemble method assumes that the individual members represent an equally likely future reality, where the spread of the ensemble is used to infer uncertainty in the projections. However, some models share the same underlying code, so a low ensemble spread cannot guarantee the realism of the projections. Some modelers have even argued that a multi-model application is not a beauty contest and that an unweighted multi-model mean is the best approach (Christensen et al., 2010). Rather, within Paper II, we argue that the evaluation of models over the historical period is an important step to establish their reliability and in some cases, the use of a sub-selection of models may be preferable when creating hydrological projections. Some authors have used correlation (e.g. Wilcke et al., 2013; Li et al., 2014) to characterize the relationship between variables over the historical period. However, some hydrological processes do not follow a linear pattern. Snowmelt for instance is a threshold-dependent process, where simulated accuracy around 0°C is especially important. To overcome the limitations associated with standard statistical evaluation tools, we proposed to use a more process-based investigation of climate model simulations based on a modeling framework that captures the intervariable relationship between temperature and precipitation. The realism of simulated streamflow is a particularly good indicator of model and bias correction performance, because streamflow is controlled by a wide range of hydrometeorological processes.

The value of ranking and choosing climate models based on hydrological performance is dependent on the end goal of the study. Ranking raw GCM-RCMs according to their hydrological performance can provide new insight for climate model developers and users. By utilizing a wide range of discharge metrics as a way to rank the performance of climate

models, this provides an automatic way to assess the interactions between atmospheric variables. However, raw GCM-RCMs were shown to have considerable biases, and should not be used for impact modeling. On the other hand, it is important to underline that QM improves the simulated discharge from a hydrological model however, it also causes convergence of the hydrological simulations. This is due to the use of a common reference observational dataset for the bias correction of the simulated atmospheric forcing. During QM, the CDFs of the GCM-RCM temperature and precipitation are forced to match that of observation. Therefore, after bias correction, the differences between the climate models is less discernible. However, if the end-goal of a given study is to perform exceptionally well on certain metrics or if intervariable relationships are exceedingly important, the small differences between the hydrological performance of the bias corrected GCM-RCMs may be of interest. In such a circumstance, ranking bias corrected GCM-RCMs would be preferable. In addition, the application of this ranking depends on the scale of the study. A well validated individual catchment may be sufficient to extend this ranking into the future and base model selection upon. However, if the study covers larger regions with, for instance, ungauged basins, the standards for culling models should be adjusted accordingly to allow for a greater uncertainty in the system being modeled (Krysanova et al., 2018).

## 5.2 Sensitivity of the modeling chain

A sensitivity assessment can be performed by making isolated adjustments to aspects of the modeling chain design, while keeping the other components of the chain the same. The resultant perturbation within the simulations is used to quantify the sensitivity of the system to the changes made. Generally speaking, sensitivity tests are performed in order to identify which areas of the modeling chain require more attention and thereby determine whether a modeler should invest additional effort in improving these aspects of the modeling chain. In this dissertation, we analyzed system sensitivity to changes in (i) bias correction (used to adjust the representation of intervariable relationships) and (ii) $E_{pot}$ estimation.

Regarding point (i) observed intervariable dependencies are often misrepresented in climate model simulations and these biases can be retained when using a univariate bias correction method. Given this limitation, it is useful to compare the impact of a univariate versus a multivariate bias correction method in order to assess the sensitivity of the modeling chain to the representation of intervariable relationships. Biases in the interdependency between temperature and precipitation were found within Paper III for raw climate output from 10 GCM-RCMs when compared to a historical observational dataset. In snow-dominated environments, the representation of atmospheric intervariable relationships is especially important, given that these areas frequently fluctuate around the threshold temperature of 0°C, which determines whether precipitation will fall as either

rain or snow. As air temperature determines the distinction between rain and snow, differences in the representation of atmospheric intervariable relationships can lead to differences in snow accumulation and therefore water storage. For MBCn-corrected data in Paper III, there was discernibly more precipitation at air temperatures below 0°C in comparison to the QDM-corrected data. This difference in the representation of intervariable relationships resulted in more snow accumulation for the MBCn driven simulations than for the QDM driven simulations. Glacier dynamics were also influenced by the choice of bias correction method. Within MBCn driven simulations, the greater snow accumulation impacted the glaciers, with higher winter mass balances and a later start of the melt season. Together, these factors directly influence streamflow composition and the evolution of the regime.

In regard to point (ii), current research on evaporation and observed continental dryness trends show contradicting futures (Greve et al., 2014; Sheffield et al., 2012). Two prominent schools of thought have emerged within the last decade. The first school of thought can be seen in work such as Oudin et al. (2005), which shows that conceptual hydrological models are often not capable of utilizing detailed information for calculating $E_{pot}$. Within hydrological modeling literature, it is common to use mean $E_{pot}$ (i.e. the same seasonally variable $E_{pot}$ for every year) instead of using temporally varying $E_{pot}$ as input to hydrological models. In addition, Andréassian et al. (2004) showed that hydrological models are clearly sensitive to $E_{pot}$ input, but use other parameters to adapt to various $E_{pot}$ scenarios. These studies suggest that data intensive $E_{pot}$ equations (e.g. Penman) require more data and effort than can even be utilized within climate change impact studies. Instead simple temperature-based $E_{pot}$ approaches, which only require mean air temperature (derived from long-term averages), lead to a slight improvement in model efficiency. The other school of thought is described nicely in Sheffield et al. (2012), which found that using simplified temperature-based methods for calculating $E_{pot}$ would likely result in an overestimation of drought. Instead Sheffield proposed that temperature-based methods for calculating $E_{pot}$ do not perform well in climate change studies. In practice, data is often a limiting factor as to which $E_{pot}$ formula is used.

When considering the data availability issues of Northwest Africa, it is therefore not surprising that hydrological projection studies performed over this region have tended to utilize temperature-based $E_{pot}$ formulas (e.g., Marchane et al., 2017; Ruelland et al., 2015). Given the dominance of evaporation within the water balance of arid regions and the issue of data availability (often resulting in the use of temperature-based formulas), we tested the impact of $E_{pot}$ estimation on hydrological projections for Northern Tunisia in Paper IV. By decomposing the variance of the projections using an ANOVA test (see Supplementary Materials of Paper IV), we found that GCM-RCMs are the dominant source of uncertainty and $E_{pot}$ estimation represents only a small proportion of the total variance. This means that it is relatively unimportant which $E_{pot}$ formulation is used within a hydrological climate change impact assessment for this type of region. Therefore, a simple, temperature-based $E_{pot}$ formulation can be considered sufficient. We also found that the

hydrological response of the Northern Tunisian catchments showed a high sensitivity to any change in precipitation, referred to as 'elasticity' (Chiew, 2006). An analysis of the elasticity of discharge to precipitation showed that for every unit change of precipitation, discharge can be expected to change by a factor of two. Given that Tunisia already suffers from water shortage issues, these results suggest this region could face extreme water stress in the future.

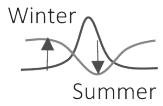
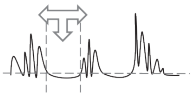
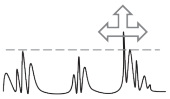
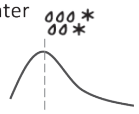
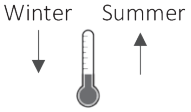
## 5.3 Application of modeling chain to aid decision making

One of the objectives of this thesis was to apply a bottom-up climate change impact modeling chain to a real-world decision making problem. By following a bottom-up approach within Paper V, our study revealed that indices often chosen by impact modelers are not necessarily fit for purpose and should be tailored to support decision making. Through project meetings with Groupe E representatives, it was found that commonly known indices such as the long-term mean monthly distribution of inflow are useful, however the complexity of their concession negotiations warranted the selection of some lesser known indices. For example, the index of 'consecutive days of low flows' was specifically requested, which typically doesn't appear in top-down climate change impact studies. This index was selected by Groupe E because their operations are vulnerable to extended periods of low flows and the development of low flow behavior in the future will likely influence the price they pay for their water fee. Additionally, Groupe E pointed out that a singular index or non tailored indices are of limited use.

To anticipate climate change impacts and develop adaptation strategies, it is important to first understand an end-users vulnerabilities. Top-down studies typically provide an overview of the impacts of climate change on hydrological resources, yet, water managers require more specific and local information in order to make informed decisions (Vano et al., 2018). The work within Paper V resulted in key findings that were directly useful to Groupe E. For instance, projections indicate a significant increase of inflow over the winter period ($\backsim$ +60% under RCP 4.5 and $\backsim$ +90% under RCP 8.5), coinciding to when energy prices have historically been at their highest. An increase of inflow over the winter period could result in profitable opportunities for Groupe E. In the coming years, Groupe E may want to analyze their ability to capture and turbinate the projected increase of winter inflow and invest in their trading and forecasting departments. Regulations allow them to sell its electricity up to 3 years in advance, hence they will need to find a balance between best price and risk management. Results also show areas where Groupe E may face challenges in the future. Projections show a reduction of summer inflows ($\backsim$ -30% under RCP 4.5 and $\backsim$ -50% under RCP 8.5) for both catchments. A reduction of inflow over the summer could put Groupe E at risk of not meeting their minimum flow requirements. This stressful circumstance may be exacerbated by a simultaneous increase in energy demand for cooling due to increasing air temperature. Low flows are expected to decrease by over -50% over JJA and SON and consecutive days of low flow will increasingly extend

beyond Groupe E's 60-day threshold. The flexibility of hydropower operations to release water when prices are optimal is a major component of what makes reservoirs profitable. Extended periods of low flow will reduce Groupe E's ability to turbinate water at optimal times, and thus lower their ability to take advantage of peak prices. Groupe E stated our projections of low flows could be used to provide a basis from which Groupe E may negotiate a lower cost for their water fee, which is partially determined based on the amount of water available for electricity production. Table 5.1 provides a summary of the opportunities and risks for Groupe E's operations in light of climate change.

TABLE 5.1: The major opportunities and risks for Groupe E's operations summarized in relation to the hydrological and climatological considerations for concession renewal. Table from Paper V.

| | | | | | |
|---|---|---|---|---|---|
| Vulnerability | Overall decrease in annual inflows and seasonal change in the distribution of inflow | Increase in the duration of low flows while simultaneously less water is carried by low flows | Seasonal change in the behaviour of high flows | Meltwater mixing with rain events while reservoir levels approach annual maximum | Electricity demand – decrease over winter and increase over summer |
| Index | • Seasonal means | • Q5<br>• Consecutive days below Q5 | • Q95<br>• Consecutive days above Q95 | • Rain versus snow contribution to runoff | • HDD<br>• CDD |
| Change in index | Increase in long-term mean monthly inflow over winter<br><br>Decrease in long-term mean monhtly inflow over summer and fall<br>  | Reduction of the magnitude of water carried by Q5 by 50% over summer and fall<br><br>Duration of low flows below Q5 will likely extend as long as 80-90 days consecutive<br>  | Duration of high flows above Q95 will likely extend as long as 20 consecutive days over winter<br><br>Likely decrease in summer high flows<br>  | Peak annual contribution from snowpack will likely shift from May to April<br><br>Rain will likely increase its contribution to inflow during the winter<br>  | Likely decrease in demand for electricity over winter<br><br>Likely increase in demand for electricity over summer and fall<br>  |
| Opportunity – adaptation option | <u>Winter</u><br>• Increase in inflows when electricity prices have been historically high could result in profit<br><br><u>Summer</u><br>• Summer inflows more unreliable<br>• Reduced inflows will likely make it harder to meet reservior level requirements | <u>Summer</u><br>• Less water & longer periods of low flows will likely make it harder to meet residual flow requirements<br>• Negotiate new/flexible residual flow requirements as part of concession<br>• These projections could provide a basis for price reduction for water fee | <u>Winter</u><br>• Increase in inflows when prices are historically high could result in profit<br><br><u>Summer</u><br>• High flows are likely to be unreliable to offset extended periods of low flows<br>• Diversify electricity mix | <u>Spring</u><br>• Fast runoff when reservoir levels are high could make reservoir level management precarious<br><br><u>All seasons</u><br>• High intensity events could be turned to profit if forecasted early – lock in trade/sale based on forecast & bolster forecast systems | <u>All seasons</u><br>• Future electricity demand is highly uncertain<br>• Suggest flexible concession terms<br>• A potential increase in dependency on renewables likely means intermittency in production, which could represent a challenge in meeting future demand |

# Chapter 6

# Conclusions

This dissertation includes a collection of studies, which together are meant to progress the numerical modeling approach to achieve projections of climate change impacts on water resources, as well as to support climate change adaptation. Specifically, the goals of this thesis were to understand and summarize the individual components of a typical impact modeling chain (Paper I), rethink evaluation methods of climate models and their bias correction (Paper II), test system sensitivity when we adjust parts of the modeling chain (Papers III & IV), and apply a climate change impact modeling chain to a real-world decision making problem (Paper V). The main conclusions of this dissertation can be summarized as follows:

**Guidance on hydrological modeling of climate change impacts:** A summary of the working methodology to carry out such an impact assessment was previously missing from the literature. In Paper I, we provide relevant information to understand the uncertainties, best practices, and common mistakes related to climate change impact modeling. Supplementary materials were also created to provide guidance on some of the key tasks within a typical modeling chain. As the title of this dissertation suggests, we aimed to provide guidance on the modeling chain structure in order to increase its application. Increasing the application of the hydrological climate change impact modeling chain is both necessary and timely given the current and impending threats imposed by climate change.

**Hydrological modeling for evaluating climate models and their bias correction:** Paper II demonstrated an evaluation of climate models and their bias correction based on hydrological performance and therefore intervariable relationships. Intervariable relationships have been overlooked within typical evaluation techniques, which is a concern given that climate change impacts are the result of interactions between variables. Quantile mapping was found to be very effective, improving discharge metrics with a 91% success rate. The performance of discharge simulations was directly traceable to the bias correction treatment and the associated improvements to the simulations of meteorological variables and their interaction with one another. This analysis was performed over the historical period, which is useful but nonetheless limited because this evaluation technique does not account for processes that could occur in a changed climate. A ranking of climate models was provided based on their ability to replicate a diverse selection of discharge metrics. This ranking is meant to provide guidance on which models are most capable

of simulating the intervariable relationships necessary for realistic discharge simulations. This ranking and methodology provides new insights to climate modelers and end users and represents a novel way to assess the realism and support the selection of climate models for climate change impact studies.

**Sensitivity evaluation of modeling chain to bias correction method and $E_{pot}$ estimation:** Papers III & IV systematically tested the sensitivity of a climate change impact modeling chain to different methods of bias correction and the use of varying levels of $E_{pot}$ information used as input. These studies convey messages that are useful beyond the specific case studies presented. Specifically, the effect of different bias correction methods on the projections of subcomponents of flow shows that the interdependence of climate variables should be taken into consideration in climate change impact studies. Conversely, our results demonstrate that projections of discharge in semi-arid regions are not sensitive to $E_{pot}$ estimation since $E_{act}$ is mainly controlled by the availability of soil moisture. Results also showed that discharge in an arid region such as Northern Tunisia has a high sensitivity to changes in precipitation. For every unit change of precipitation, discharge was found to change by a factor of two. Given this region's propensity towards water shortages, these projections signal the need for additional studies on this topic.

**Application of modeling chain to aid decision-making:** Paper V built upon the concept that an end-users' needs and vulnerabilities should be well understood in order to provide projections useful for decision making. By tailoring our modeling chain to the needs of the end-user, this allowed for the vulnerabilities and opportunities of the hydropower company to be expressed within the figures. Our results were visualized in a way that provided support to negotiate the renewal of the hydropower company's water concession. They stated that they were interested in conducting similar studies for their other catchments and would consider investing in new hydrological projections at some point in the future. They also stated that this collaboration made the climate change phenomenon much more real and expressed that our collaboration was useful in allowing them to envision the impacts they are likely to experience under the influence of climate change.

This research is especially timely given the multi-decade length of a water concession, wherein the environmental regulations cannot be further adjusted once agreed upon. Therefore, projections of future water availability are especially useful in helping to determine whether the environmental regulations proposed today will be realistically attainable in the coming decades. The results of this study serve a wider community than just a single hydropower company. Across Switzerland, the vast majority of all water concessions are still to be negotiated in the coming years. This study provides guidance on how available research methods can be modified and used to support the adaptation of the hydropower industry to climate change.

# Chapter 7

# Future Research

Although many topics discussed within this dissertation could be expanded, the aim of this dissertation is to develop the numerical modeling chain to increase its application and to aid in climate change adaptation and decision support. Through the process of conducting each research study, some topics were found to offer opportunities to address this goal more directly than others. Future research projects are discussed below with this in mind:

**$E_{pot}$ estimation within climate change impact studies:** Paper IV of this dissertation offers significant potential for expansion, given that a future change in evaporation is of special concern due to its influence on the water-energy cycle. Within the scientific literature, there has been an expectation that the hydrological cycle will accelerate under climate change (IPCC, 2013). As global air temperature increases and with more energy for evaporation, it was assumed that this would have resulted in an increase in atmospheric evaporative demand. However the opposite has been observed in pan evaporation measurements (Johns Hopkins University, 1998). This discovery by Peterson et al. (1995) is now known as the pan evaporation paradox, and has since been confirmed by observations worldwide (Brutsaert and Parlange, 1998; Zhang et al., 2016; Roderick and Farquhar, 2002; Hobbins et al., 2004). One overlooked aspect is whether the evaporation paradox would appear, and what would be the underlying causes in the future, within GCM simulations (Wang et al., 2017). The challenge still remains to reconcile past evaporative demand with that under future climate. Further still, it is important to relate changes in evaporative demand to the rest of the water cycle in a realistic way. As Paper IV established, the amount of $E_{pot}$ information provided to hydrological models is essentially irrelevant within an arid region, given that these areas are water limited. However, Paper IV only evaluated mean flows and all catchments were located within the same climatic zone. Contrasting results were found in a preliminary study for this dissertation, performed by Claudia Teutschbein. In the preliminary study, two different setups were tested according to their ability to simulate streamflow in the Vattholmaan catchment located in southeastern Sweden. Within the first setup, bias-corrected daily temperature and precipitation simulations from six RCMs were downloaded from the ENSEMBLES (predecessor to CORDEX) web page. $E_{pot}$ was calculated based on long-term means and used as input to HBV. In the second setup, bias-corrected daily $E_{pot}$ was taken as direct output from the same six RCMs and used to

drive HBV. Results showed distinctly different hydrological simulations between the two setups, especially in projections of the length of future low flows and the calculated annual 15-day low flow (statistic based on an annual series of the smallest values of mean discharge computed over any 15-consecutive day period). These differences illustrate the need to understand how much uncertainty is introduced in discharge projections by the formulation of $E_{pot}$, especially for indices besides mean flow and in non-arid environments.

**Multidisciplinary approach to support water management:** Paper V represents a research area that requires immediate attention. Hydropower managers are currently faced with the challenges spurred on by the ending of their water concessions. Now and in the coming decades, these managers must decide whether or not to renew their water concession, invest in new infrastructure, purchase new shares in additional reservoirs or invest in different energy sources.

One step forward would be to explicitly model hydropower operations as done in Anghileri et al. (2018) and Finger et al. (2012), which could bolster the usefulness of any projections created. To model hydropower operations, Finger et al. (2012) utilized a model that accounts for storage in reservoirs, water abstraction, diversion of natural mountain streams, and pressurized water to turbines. To this extent, the work shown in Paper V could be expanded by increasing modeling complexity to incorporate the behavior of the reservoir operations.

In addition, a collaboration between hydrologists, economists, and stakeholders would be beneficial to ensure that all necessary components are included within a newly designed modeling framework. An important aspect missing from Paper V is the lack of consideration for water or electricity demand within the modeling framework, although the indices of $C_{DD}$ and $H_{DD}$ were used to superficially explore this topic. Other studies that have tried to address this topic have fallen upon two recurring issues. First, studies that provide long-term projections of electricity prices (Savelsberg et al., 2018) or future energy mix (Redondo and Van Vliet, 2015), do not consider a wide enough selection of hydrological indices relevant for hydropower management (i.e. they tend to only focus on changes to long-term mean monthly inflow to the reservoirs). Secondly, the electricity market is extremely volatile, and therefore expanding Paper V to incorporate electricity price scenarios poses many limitations. An initial effort was made by Anghileri et al. (2018) to combine hydrological and electricity market modeling, to assess the sensitivity of an Alpine hydropower system to changes in water availability and electricity price. However, this study doesn't explicitly discuss hydrological changes and their relation to hydropower operations, but rather aggregates these impacts by focusing on the profitability of the reservoir. This is informative but unfortunately the aggregation of hydrological impacts reduces the information that could be used for climate change adaptation efforts. A pragmatic step forward would be to retain the information gleaned from individual streamflow indices and expand upon those chosen within Paper V. In particular, indices related to the magnitude and duration of flooding would be especially useful. Indices could also be evaluated

according to their universal appropriateness for hydropower climate change adaptation.

**Scenario neutral approach to support water management:** Given that there are many unknowns regarding our understanding of how the Swiss and European electricity markets will develop, and that there are uncertainties related to climate change impacts on hydrological regimes, an alternative modeling framework could be used to progress the efforts of Paper V. The scenario-neutral approach, developed by Prudhomme et al. (2010), provides insight to a system's response to a range of climatic changes. Similar to elasticity, the scenario-neutral approach can capture nonlinear relationships between variables. It involves the testing of the responsiveness of a local indicator (e.g. reservoir) to incremental adjustments of a driving variable (e.g. temperature, precipitation). Although these incremental adjustments are not necessarily related to expected future climate, these methods could be further developed, and these studies should be commended for their apparent efforts to shift focus to end-user needs. Scenario-neutral approaches have been conducted for flood risk management (e.g. Broderick et al., 2019), hydrological indices (e.g. Singh et al., 2014), and urban water supply (e.g. Brown and Wilby, 2012). Most relevant to this dissertation, is the work performed by Singh et al. (2014), which utilized a scenario-neutral approach by first asking end-users to define the vulnerable ranges (i.e. thresholds) for hydrological indices. The modeled system was then analyzed to assess under which conditions these thresholds are exceeded. Future work could combine aspects of the scenario-neutral approach with the modeling chain described in this dissertation in order to aid water managers with climate change adaptation efforts.

# References

Alexander, L. V. et al. (2006). "Global observed changes in daily climate extremes of temperature and precipitation". *Journal of Geophysical Research* 111, pp. 1–22. DOI: 10.1029/2005JD006290.

Andréassian, V., C. Perrin, and C. Michel (2004). "Impact of imperfect potential evapotranspiration knowledge on the efficiency and parameters of watershed models". *Journal of Hydrology* 286, pp. 19–35. DOI: 10.1016/j.jhydrol.2003.09.030.

Anghileri, D. et al. (2018). "A Comparative Assessment of the Impact of Climate Change and Energy Policies on Alpine Hydropower". *Water Resources Research* 54, pp. 9144–9161. DOI: 10.1029/2017WR022289.

Arora, V. K. and G. J. Boer (2001). "Effects of simulated climate change on the hydrology of major river basins". *Journal of Geophysical Research* 106, pp. 3335–3348. DOI: 10.1029/2000JD900620.

Barry, M. et al. (2015). "The Future of Swiss Hydropower A Review on Drivers and Uncertainties". *SCCER CREST Workpackage 3: Energy Policy, Markets and Regulation*, pp. 1–49.

Bartholomeus, R. P. et al. (2015). "Sensitivity of potential evaporation estimates to 100 years of climate variability". *Hydrology and Earth System Sciences* 19, pp. 997–1014. DOI: 10.5194/hess-19-997-2015.

Bergström, S. (1976). *Development and Application of a Conceptual Runoff Model for Scandinavian Catchments*. Vol. RHO 7. Norrköping, p. 134.

Bergstrom, S. (1992). *The HBV model*. Tech. rep. 4.

Breuer, L. et al. (2009). "Assessing the impact of land use change on hydrology by ensemble modeling (LUCHEM). I: Model intercomparison with current land use". *Advances in Water Resources* 32, pp. 129–146. DOI: 10.1016/j.advwatres.2008.10.003.

Broderick, C. et al. (2019). "Using a Scenario-Neutral Framework to Avoid Potential Maladaptation to Future Flood Risk". *Water Resources Research*, pp. 1079–1104. DOI: 10.1029/2018WR023623.

Brown, C. and R. L. R. L. Wilby (2012). "An alternate approach to assessing climate risks". *Eos, Transactions American Geophysical Union* 93, pp. 401–402. DOI: 10.1029/2012EO410001.

Brutsaert, W. and M. B. Parlange (1998). "Hydrologic cycle explains the evaporation paradox". *Nature* 396, p. 30. DOI: 10.1038/23845.

Cannon, A. J. (2018). "Multivariate quantile mapping bias correction: an N-dimensional probability density function transform for climate model simulations of multiple variables". *Climate Dynamics* 50, pp. 31–49. DOI: 10.1007/s00382-017-3580-6.

Chen, J. et al. (2013). "Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America". *Water Resources Research* 49, pp. 4187–4205. DOI: 10.1002/wrcr.20331.

Chiew, F. H. S. (2006). "Estimation of rainfall elasticity of streamflow in Australia". *Hydrological Sciences Journal* 51, pp. 613–625. DOI: 10.1623/hysj.51.4.613.

Christensen, J. H. et al. (2010). "Weight assignment in regional climate models". *Climate Research* 44, pp. 179–194. DOI: 10.3354/cr00916.

Dickinson, R. E. et al. (1989). "A Regional Climate Model for the Western United States". *Climatic Change* 15, pp. 383–422.

Finger, D. et al. (2012). "Projections of future water resources and their uncertainty in a glacierized catchment in the Swiss Alps and the subsequent effects on hydropower production during the 21st century". *Water Resources Research* 48, pp. 1–20. DOI: 10.1029/2011WR010733.

Frei, C. (2013). "Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances". *International Journal of Climatology* 34, pp. 1585–1605. DOI: 10.1002/joc.3786.

Frei, C. and C. Schär (1998). "A precipitation climatology of the Alps from high-resolution rain-gauge observations". *International Journal of Climatology* 18, pp. 873–900. DOI: 10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9.

Frick, C. et al. (2014). "Central European high-resolution gridded daily data sets (HYRAS): Mean temperature and relative humidity". *Meteorologische Zeitschrift* 23, pp. 15–32. DOI: 10.1127/0941-2948/2014/0560.

Gaudard, L., M. Gilli, and F. Romerio (2013). "Climate Change Impacts on Hydropower Management". *Water Resources Management* 27, pp. 5143–5156. DOI: 10.1007/s11269-013-0458-1.

Greve, P. et al. (2014). "Global assessment of trends in wetting and drying over land". *Nature Geoscience* 4, pp. 1–6. DOI: 10.1038/ngeo2247.

Gudmundsson, A. L. (2016). *Statistical Transformations for Post-Processing Climate Model Output*. Tech. rep. DOI: 10.5194/hess-16-3383-2012.bernexp.

Gudmundsson, L. et al. (2012). "Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations - a comparison of methods". *Hydrology and Earth System Sciences* 16, pp. 3383–3390. DOI: 10.5194/hess-16-3383-2012.

Hirabayashi, Y. et al. (2013). "Global flood risk under climate change". *Nature Climate Change* 3, pp. 816–821. DOI: 10.1038/nclimate1911.

Hobbins, M. T., J. A. Ramírez, and T. C. Brown (2004). "Trends in pan evaporation and actual evapotranspiration across the conterminous U.S.: Paradoxical or complementary?" *Geophysical Research Letters* 31, pp. 1–5. DOI: 10.1029/2004GL019846.

Hrachowitz, M. and M. P. Clark (2017). "HESS Opinions : The complementary merits of competing modelling philosophies in hydrology". *Hydrology and Earth System Sciences*, pp. 3953–3973. DOI: 10.5194/hess-21-3953-2017.

Huntington, T. G. (2006). "Evidence for intensification of the global water cycle: Review and synthesis". *Journal of Hydrology* 319, pp. 83–95. DOI: 10.1016/j.jhydrol.2005.07.003.

IPCC (2001). *Climate change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the IPCC*. Ed. by J. Houghton et al. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, p. 881.

— (2007). *Towards new scenarios for analysis of emissions, climate change, impacts, and response strategies*.

— (2013). *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by T. F. Stocker et al. Cambridge, United Kingdom and New York, NY, USA, p. 1535.

— (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Ed. by Core Writing Team, R. Pachauri, and L. Meyer. Geneva, Switzerland: IPCC, Geneva, Switzerland, p. 151. DOI: 10.1017/CBO9781107415324. arXiv: arXiv:1011.1669v3.

Jakeman, A., I. Littlewood, and P. Whitehead (1990). "Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments". *Journal of Hydrology* 117, pp. 275–300. DOI: 10.1016/0022-1694(90)90097-H.

Johns Hopkins University (1998). *Environmental Engineers Unravel Evaportation Paradox*.

Knutti, R. and J. Sedláček (2013). "Robustness and uncertainties in the new CMIP5 climate model projections". *Nature Climate Change* 3, pp. 369–373. DOI: 10.1038/nclimate1716.

Kotlarski, S. et al. (2014). "Regional climate modeling on European scales: A joint standard evaluation of the EURO-CORDEX RCM ensemble". *Geoscientific Model Development* 7, pp. 1297–1333. DOI: 10.5194/gmd-7-1297-2014.

Ková, P. et al. (2015). "Choosing an appropriatehydrological model for rainfall-runoff extremes in small catchments". *Soil and Water Resources* 2015.3, pp. 137–146. DOI: 10.17221/16/2015-SWR.

Krause, P and D. P. Boyle (2005). "Comparison of different efficiency criteria for hydrological model assessment". *Advances In Geosciences* 5, pp. 89–97. DOI: 10.5194/adgeo-5-89-2005.

Krysanova, V. et al. (2018). "How the performance of hydrological models relates to credibility of projections under climate change". *Hydrological Sciences Journal* 63, pp. 696–720. DOI: 10.1080/02626667.2018.1446214.

Lee, J. (1993). "A formal approach to hydrological model conceptualization". *Hydrological Sciences Journal* 38:5, pp. 391–401. DOI: 10.1080/026266693099492689.

Li, C. et al. (2014). "Joint bias correction of temperature and precipitation in climate model simulations". *Journal of Geophysical Research : Atmospheres* 119.13, pp. 153–162. DOI: 10.1002/2014JD022514.

Lindström, G. et al. (1997). "Development and test of the distributed HBV-96 hydrological model". *Journal of Hydrology* 201, pp. 272–288. DOI: 10.1016/S0022-1694(97)00041-3.

Manabe, S. and R. T. Wetherald (1967). "Thermal equilibrium of the atmosphere with a given distrution of relative humidity". *Journal of the Atmospheric Sciences* 24, pp. 241–259.

Maraun, D and M Widmann (2018). *Statistical Downscaling and Bias Correction for Climate Research*. Cambridge University Press, Cambridge, UK. DOI: 10.1017/9781107588783.

Maraun, D. (2016). "Bias Correcting Climate Change Simulations - a Critical Review". *Current Climate Change Reports* 2, pp. 211–220. DOI: 10.1007/s40641-016-0050-x.

Marchane, A. et al. (2017). "Climate change impacts on surface water resources in the Rheraya catchment (High Atlas, Morocco)". *Hydrological Sciences Journal* 62, pp. 979–995. DOI: 10.1080/02626667.2017.1283042.

Mauch, C. and E. Reynard (2004). *The Evolution of National Water Regimes in Switzerland*. DOI: 10.1007/978-1-4020-2484-9_9.

McMahon, T. A., B. L. Finlayson, and M. C. Peel (2016). "Historical developments of models for estimating evaporation using standard meteorological data". *WIRES Water* 3, pp. 788–818. DOI: 10.1002/wat2.1172.

Mcguffie, K. and A. Henderson-Sellers (2001). "Forty years of numerical climate modelling". 1109.2001, pp. 1067–1109.

MeteoSwiss (2013). *Documentation of MeteoSwiss grid-data products: Daily precipitation (final analysis): RhiresD*. Tech. rep. Federal Office of Meteorology and Climatology MeteoSwiss, pp. 1–4. DOI: OFEV2014.

Nash, J. and J. Sutcliffe (1970). "River flow forecasting through conceptual models part I A discussion of principles". *Journal of Hydrology* 10, pp. 282–290. DOI: 10.1016/0022-1694(70)90255-6.

Oudin, L., C. Michel, and F. Anctil (2005). "Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 1 - Can rainfall-runoff models effectively handle detailed potential evapotranspiration inputs?" *Journal of Hydrology* 303, pp. 275–289. DOI: 10.1016/j.jhydrol.2004.08.025.

Parry, M., J. Lowe, and C. Hanson (2009). "Overshoot, adapt and recover". *Nature Climate Change* 458, pp. 1102–1103. DOI: 10.1038/climate.2008.50.

Perrin, C., C. Michel, and V. Andréassian (2003). "Improvement of a parsimonious model for streamflow simulation". *Journal of Hydrology* 279, pp. 275–289. DOI: 10.1016/S0022-1694(03)00225-7.

Peterson, T. C., V. S. Golubev, and P. Y. Groisman (1995). "Evaporation is loosing its strength". *Nature* 377, pp. 687–688.

Prudhomme, C. and J. Williamson (2013). "Derivation of RCM-driven potential evapotranspiration for hydrological climate change impact analysis in Great Britain: A comparison of methods and associated uncertainty in future projections". *Hydrology and Earth System Sciences* 17, pp. 1365–1377. DOI: 10.5194/hess-17-1365-2013.

Prudhomme, C. et al. (2010). "Scenario-neutral approach to climate change impact studies: Application to flood risk". *Journal of Hydrology* 390, pp. 198–209. DOI: 10.1016/j.jhydrol.2010.06.043.

Rauthe, M. et al. (2013). "A Central European precipitation climatology - Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS)". *Meteorologische Zeitschrift* 22, pp. 235–256. DOI: 10.1127/0941-2948/2013/0436.

Redondo, P. D. and O. Van Vliet (2015). "Modelling the energy future of Switzerland after the phase out of nuclear power plants". *Energy Procedia* 76, pp. 49–58. DOI: 10.1016/j.egypro.2015.07.843.

Roderick, M. L. and G. D. Farquhar (2002). "The cause of decreased pan evaporation over the past 50 years". *Science* 298, pp. 1410–1411. DOI: 10.1126/science.1075390.

Ruelland, D., P. Hublart, and Y. Tramblay (2015). "Assessing uncertainties in climate change impacts on runoff in Western Mediterranean basins". *IAHS-AISH Proceedings and Reports* 371, pp. 75–81. DOI: 10.5194/piahs-371-75-2015.

Savelsberg, J. et al. (2018). "The impact of climate change on Swiss hydropower". *Sustainability* 10, pp. 1–23. DOI: 10.3390/su10072541.

Schwarb, M. (2000). "The alpine precipitation climate: Evaluation of a high-resolution analysis scheme using comprehensive rain-gauge data". PhD thesis. Swiss Federal Institute of Technology Zurich, p. 131. DOI: 10.3929/ethz-a-004121274.

Seibert, J. and M. J. P. Vis (2012). "Teaching hydrological modeling with a user-friendly catchment-runoff-model software package". *Hydrology and Earth System Sciences* 16, pp. 3315–3325. DOI: 10.5194/hess-16-3315-2012.

Sheffield, J., E. F. Wood, and M. L. Roderick (2012). "Little change in global drought over the past 60 years." *Nature* 491, pp. 435–8. DOI: 10.1038/nature11575.

Singh, R. et al. (2014). "A vulnerability drive approach to identify adverse climate and land use change combinations for critical hydrologic indicator thresholds: Application to a watershed in Pennsylvania, USA". *Water Resources Research* 50, pp. 1–17. DOI: 10.1002/2013WR014988.

Stahl, K. et al. (2010). "Streamflow trends in Europe: evidence from a dataset of near-natural catchments". *Hydrology and Earth System Science* 14, pp. 2367–2382. DOI: 10.5194/hess-14-2367-2010.

Teutschbein, C. and J. Seibert (2012). "Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods". *Journal of Hydrology* 456-457, pp. 12–29. DOI: 10.1016/j.jhydrol.2012.05.052.

Thompson, J. R., A. J. Green, and D. G. Kingston (2014). "Potential evapotranspiration-related uncertainty in climate change impacts on river flow : An assessment for the Mekong River basin". *JOURNAL OF HYDROLOGY* 510, pp. 259–279. DOI: 10.1016/j.jhydrol.2013.12.010.

Todini, E. (2011). "History and perspectives of hydrological catchment modelling". *Hydrology Research* 42, p. 73. DOI: 10.2166/nh.2011.096.

Todini, E. (2007). "Hydrological catchment modelling: past, present and future". *Hydrology and Earth System Sciences* 11, pp. 468–482. DOI: 10.5194/hess-11-468-2007.

Tonka, L. (2015). "Hydropower license renewal and environmental protection policies: a comparison between Switzerland and the USA". *Regional Environmental Change* 15, pp. 539–548. DOI: 10.1007/s10113-014-0598-8.

UNDP (2007). *Human Development Report 2007/2008. Fighting climate change: Human solidarity in a divided world.*

UNEP (2015). *Options for Decoupling Economic Growth from Water use and Water Pollution.* DOI: 10.18356/d38f0de2-en.

Vano, J. A. et al. (2018). "DOs and DON'Ts for using climate change information for water resource planning and management: guidelines for study design". *Climate Services* 12, pp. 1–13. DOI: 10.1016/j.cliser.2018.07.002.

Wang, Q. J. (1991). "The Genetic Algorithm and Its Application to Calibrating Conceptual RainfallRunoff Models". *Water Resources Research* 27, pp. 2467–2471. DOI: 10.1029/91WR01305.

Wang, T. et al. (2017). "Pan evaporation paradox and evaporative demand from the past to the future over China: a review". *Wiley Interdisciplinary Reviews: Water* 4, pp. 1–13. DOI: 10.1002/wat2.1207.

Wilby, R. L. and S. Dessai (2010). "Robust adaptation to climate change". *Weather* 65, pp. 180–185. DOI: 10.1002/wea.504.

Wilcke, R. A. I., T. Mendlik, and A. Gobiet (2013). "Multi-variable error correction of regional climate models". *Climatic Change* 120, pp. 871–887. DOI: 10.1007/s10584-013-0845-x.

Zhang, Y. et al. (2016). "Multi-decadal trends in global terrestrial evapotranspiration and its components". *Scientific Reports* 6, pp. 1–12. DOI: 10.1038/srep19124.

PAPER I

# Hydrological Modeling of Climate Change Impacts

*Kirsti Hakala*[1], *Nans Addor*[2], *Claudia Teutschbein*[3], *Marc Vis*[1], *Hamouda Dakhlaoui*[4,5], *and Jan Seibert*[1,6]

[1] *University of Zurich, Zurich, Switzerland*
[2] *University of East Anglia, Norwich, UK*
[3] *Uppsala University, Uppsala, Sweden*
[4] *University of Tunis El Manar, Tunis, Tunisia*
[5] *University of Carthage, Sidi Bou Said, Tunisia*
[6] *Swedish University of Agricultural Sciences, Uppsala, Sweden*

## 1 Introduction

### 1.1 Why Use Hydrological Modeling to Project Impacts of Climate Change?

The impacts of anthropogenic climate change on the water cycle are already apparent [1–3]. These impacts include changes in annual river streamflow [4], shifts in both flood peak magnitude and timing [5], alterations in flow duration curves [6], and changes in magnitude of low-flow periods [7]. The continued increase of global temperatures will lead to further changes in regional hydrology within the next decades through shifts in precipitation trends, melting of glaciers and permafrost [8], and a growing rain-to-snow ratio in cold regions [3, 9]. In addition, changes in natural vegetation cover, land use practices, crop water requirements, prolonged growing seasons, and soil functions may further alter the hydrological cycle [10]. Extreme events such as river flooding pose a potential threat to human societies and are likely to occur more often [2, 11]. Given that these changes directly affect agriculture, forestry, energy production, drinking water supply, sanitation, and ecosystems, there are likely to be substantial consequences for societies in many regions around the world [12]. Reliable information on potential changes to future hydrological conditions is fundamental for deciding on long-term management strategies and adaptation measures [13].

Given the impact of climate change on hydrology, hydrologists are asked to provide the hydrological basis for future water development and management, which requires an understanding of the impact of climatic and environmental change on future hydrological conditions [13]. Computer models are suitable tools to obtain such quantitative information for possible future conditions. However, any model is a simplification of reality and model simulations are uncertain, especially when a combination of models is used to represent the climate and land-surface processes, as is the case in hydrological climate change impact studies. Therefore, addressing uncertainties is an important aspect of carrying out a hydrological climate change impact study.

### 1.2 Goals of the Article

This article provides relevant information to understand (i) hydrological climate change impact research, (ii) the steps to perform an impact study, and (iii) the main challenges encountered in an impact study and how they can be addressed. Hydrological climate change research is an active field of research and although much progress has been made, many challenges remain. Sometimes, these challenges are difficult to overcome and being aware of the limitations is the best one can achieve. This article does not aim to provide a complete review of all models, datasets, and methods used for hydrological climate change impact studies. Rather, we summarize the most relevant subcomponents of hydrological climate change research. Uncertainties are a main focus throughout the article and best practices to characterize them are discussed. Supplementary material includes a guide to perform some key tasks leading to the production of hydrological projections, and a basis for course material to teach the analysis of climate impacts on water resources (see Section 5.6). The materials presented here presume a working knowledge of climate and hydrological sciences,

*Encyclopedia of Water: Science, Technology, and Society,* edited by Patricia A. Maurice.

DOI: 10.1002/9781119300762.wsts0062

corresponding to approximately a bachelor's degree in, for instance, geosciences. This article was written as an introduction for researchers and students who are planning their own hydrological climate change impact assessment study. However, this article can also be used to guide experienced scientists through aspects of the modeling chain with which they are less familiar.

This article is structured into seven sections. Following the introduction, Section 2 describes the basics of the model chain as a whole, providing an overview of the standard model chain used to produce hydrological projections. Section 3 provides background information on each step of the model chain and its associated uncertainties. In Section 4 uncertainty sampling and decomposition are discussed. Section 5 provides practical guidance on how to design and carry out an impact study and introduces best practices on how to evaluate streamflow projections. Section 6 contains an overview of research questions currently addressed in the literature and makes the reader aware of limitations of the approaches presented in the previous sections. In Section 7, a general outlook of hydrological climate change impact research is provided. Important terms are explained as they appear in the text; for a glossary, please see for instance the Annex III within the Intergovernmental Panel on Climate Change's (IPCC) Working Group I Fifth Assessment Report (IPCC [14]; available at https://www.ipcc.ch/site/assets/uploads/2018/02/WG1AR5_AnnexIII_FINAL.pdf).

## 2 Overview of the Modeling Chain

To investigate hydrological behavior under climate change, projections of future climate are needed, which are generated using general circulation models (GCMs, also sometimes referred to as global climate models). Over time, GCMs have been developed to include more systems besides the ocean and atmosphere, and they now include other processes such as global carbon cycle, dynamic vegetation, or atmospheric chemistry [3]. These models now go beyond the original definition of GCM and are thus referred to as earth system models (ESM). Owing to the evolving nature of these models, there is some ambiguity amongst these terms in climate science literature. In most cases in this article, when we are referring to GCMs, we are actually discussing ESMs. However, to keep consistent with the climate modeling community, we will use the acronym GCM.

For projections of future climate, a wide range of potential scenarios is available (Figure 1), including for example greenhouse gas emission scenarios based on specific socioeconomic assumptions or so-called representative concentration pathway scenarios based on forcing projections that could, in theory, be realized with more than one socioeconomic scenario.

Because of the coarse spatial resolution of GCMs (horizontal grid spacing~100–300 km), modelers usually downscale their output to a finer resolution, using a regional climate model (RCM) or using statistical techniques. Yet, the downscaled data can still show

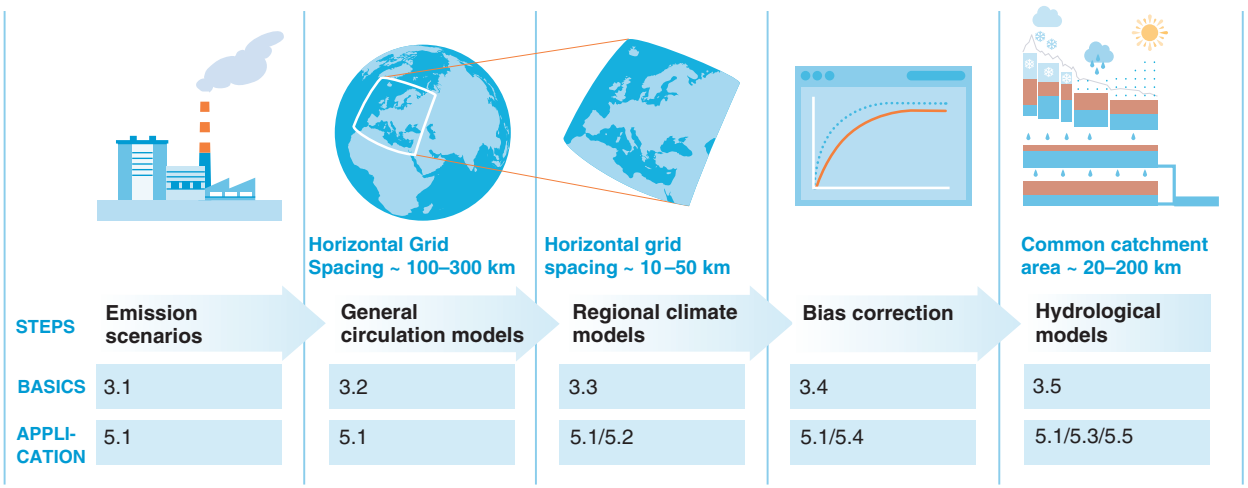



**Figure 1** Schematic of a typical model chain used for the assessment of climate change impacts on streamflow. Graphics of the modeling chain are shown in the first row with the names of the steps listed in the second row. The last two rows refer to the sections of this article in which each step is discussed.

considerable biases compared to observed data [15]. Various post-processing methods (often referred to as "bias correction") have been developed to reduce biases in simulated time series. When climate model simulations are used as input to hydrological models, such bias correction generally leads to improved discharge simulations [16]. Therefore, climate projections are typically bias corrected before being used as input to hydrological models to create projections of streamflow under the influence of different climate scenarios (Figure 1). The calibration of the hydrological model (usually using observational meteorological and hydrological data) is performed in order to tune the parameters of the hydrological model so that the differences between simulated and observed discharge are minimized. These optimized parameter sets are then used in the simulations where climate projections force the hydrological model. Subsequent analysis of the projected streamflow and communication of results then follow (Figure 1).

Although hydrological models are ordinarily used to generate streamflow time series, simulated discharge can also be retrieved directly from climate models. The newest generation of GCMs include detailed descriptions of the land surface (via a coupled land surface model), which make use of river routing schemes and methods (see Figure 1 within Shaad [17]). Many different river routing schemes are now available (for a review see Clark et al. [18]). Given these advancements, and depending on the catchment size, it may make sense to consider the direct hydrological changes within the GCM or RCM output. For instance Hagemann et al. [19] compared hydrological simulations from five reanalysis driven (i.e. a global historical data set describing the state of the Earth system created by combining observations with a numerical model [20]) RCMs over two large European catchments with areas of 1.8 million $km^2$ and 800 000 $km^2$. Depending on the catchment, their results showed that generally one to two of the five RCMs were capable of simulating the annual discharge cycle fairly well, however, biases were evident in all RCM derived discharge. The authors attribute these biases to systematic errors in the model dynamics or deficiencies in the land surface parameterization. Another study by González-Zeas et al. [21] analyzed discharge output from 10 RCMs forced by GCMs for mainland Spain, which has an area of 504 782 $km^2$. After applying a bias correction using an "observed" global discharge dataset [22] as a benchmark, they compared the observed annual discharge cycle with that derived using raw RCM and bias-corrected discharge. Their results show that bias-corrected discharge corresponds well to observation and the raw discharge from a few RCMs reasonably captures the annual discharge cycle. Despite such applications, the catchment sizes in impact studies are often smaller than those found within the aforementioned studies and the spatial resolution of GCM and RCM hydrological output is often not appropriate. Furthermore, biases within the atmospheric forcing will be inherited by the hydrological output of a GCM or RCM. While the use of GCM or RCM hydrological output may be worth considering for particular applications [23–25], one also has to be aware that sometimes there are large deviations of GCM or RCM simulated discharge from observations [15].

The next section discusses in more detail the main steps of the typical model chain and reviews their main uncertainties.

## 3 Steps of the Modeling Chain and Their Uncertainty

### 3.1 Emission Scenarios

Emission scenarios are based on historical greenhouse gas (GHG) concentration data and provide estimates of future GHG concentration in the atmosphere, following assumptions of how emissions will change with evolving societal elements, such as demography, economic development, energy consumption, and land use. GCMs are then run with these scenarios to create projections of the climate under changing GHG concentrations.

Emission scenarios were developed by the IPCC, established in 1988 by two United Nations organizations, the World Meteorological Organization (WMO) and the United Nations Environment Programme (UNEP). So far, the IPCC has published five assessment reports. The First Assessment Report (FAR, 1990) used SA90 scenarios and the Second Assessment Report (SAR, 1995) used the IS92a to f emission scenarios. In 2000, the Special Report on Emission Scenarios (SRES) introduced the IPCC's third generation of scenarios. The SRES scenarios were used in IPCC's Third Assessment Report (TAR) and Fourth Assessment Reports (AR4). The Fifth Assessment Report (AR5) was completed in 2014 and relies on the fourth generation of emission scenarios referred to as Representative Concentration Pathways (RCPs), which are the most comprehensive attempt to characterize global emissions so far. The RCPs feature four trajectories (RCP2.6, RCP4.5, RCP6, and RP8.5), which are named after their associated anthropogenic radiative forcing for the year 2100 (+2.6, +4.5, +6.0, and +8.5 W $m^{-2}$). The RCP2.6 scenario is the most optimistic in that it assumes that GHG concentrations peak between 2010–2020 and then decline afterward. In contrast, RCP8.5 is the most pessimistic scenario which assumes that GHG concentrations will continue

to increase throughout the twenty-first century. The Sixth Assessment Report is expected to be published in 2022.

#### 3.1.1 Uncertainties Related to Emission Scenarios

*Human component*: The uncertainties in future GHG concentration are not of the same nature as those discussed in the remainder of this article. Future emissions will be determined by political and socio-economical decisions, whereas uncertainties in other elements of the model chain stem principally from our incomplete understanding of our natural environment and its imperfect representation in models. Early on, different emissions scenarios were developed to reflect the uncertainties in future emissions. Typically, the differences between the emission scenarios remain small for the first half of the twenty-first century but they can be significant in the second half of the century [3, 26].

### 3.2 Climate Models

GCMs are computer models, which embody simplified representations of the climate system [27]. They are essential for the understanding of the Earth's global processes under past, present, and future conditions. GCMs capture the interactions between the major components of the climate system, including the atmosphere, the oceans, the biosphere, and sea ice [28]. While simpler one- or two-dimensional models have been used to provide globally or regionally averaged estimates of climate change, GCMs are based on a three-dimensional grid covering the Earth (latitude, longitude and vertical "height"), and compute atmospheric variables (such as temperature and humidity) for each grid cell. GCMs thus, have the ability to provide both spatially and physically realistic estimates of climate change, which is the first requirement for any impact analysis. For a review of climate models of different complexity and their development over time please see Kour et al. [29].

Over the years, projections from several generations of GCMs have been produced within the framework of successive Climate Model Intercomparison Projects (CMIPs), led by the World Climate Research Program (WCRP). The goal of CMIP is to coordinate the production of GCM simulations, in order to provide consistent and reliable data, used in particular for the IPCC assessment reports. With each phase of the CMIP, model projections are further improved. Knutti et al. [30] compared CMIP5, CMIP3, and CMIP2 and found that although most models are strongly related to their predecessors, the models in the new ensemble agree more closely with observations. Model output from CMIP5 were released around 2011 and constitute the reference simulations until the release of CMIP6 simulations [31]. CMIP5 data can be downloaded through a portal provided by the Earth System Grid Federation-Center for Enabling Technologies website: http://pcmdi9.llnl.gov. A "Getting started" page is available here: https://cmip.llnl.gov/cmip5/data_getting_started.html. Note that the resolution of the GCM runs is often too coarse to realistically capture processes essential for the correct representation of streamflow generation at the local scale (such as convective events or snow accumulation in mountainous areas). Hence the majority of hydrological impact assessments do not directly use GCM simulations at their original resolution, but instead, use downscaling techniques to refine the projections.

#### 3.2.1 Uncertainties Related to Climate Models

*Model structure/parameterization*: Climate models are developed by different groups across the world. These groups make different choices when deciding which processes should be represented, and how they should be represented. For instance, convection occurs at a spatial scale smaller than the grid size of GCMs, so it cannot be explicitly represented, and instead, it has to be parameterized. Different groups will use different parameterization schemes for convection, resulting in different projections of intense precipitation events. The spread among the projections of models produced by different groups reflects both limitations in our understanding of the climate system and limitations of what can be represented by climate models running at a relatively coarse resolution. To account for climate model uncertainties, it is now a standard procedure to use an ensemble of climate models, instead of a single model (see Section 4.2 for discussion of the ensemble approach to test model structural uncertainty and Section 4.3 for an ensemble approach to analyze parametric uncertainty).

*Natural variability*: The atmosphere is a chaotic system, meaning that a small perturbation in the initial conditions of a climate state can lead to large differences in the future [32]. This makes weather forecasting difficult and also reduces the predictability of future climate. For instance, Deser et al. [33] demonstrated the importance of natural climate variability by running the same climate model several times and only changing the initial conditions by introducing an infinitesimal perturbation. This led to very different trends in the projections. The spread in the projections reflects natural variability of the climatic system, which exists even in absence of climate change. Unlike other sources of uncertainty, natural variability has an inherently unpredictable nature and is unlikely to be reduced even as newer generations of climate models are unveiled. For an extended discussion of methods to analyze natural variability, see Section 4.4.

### 3.3 Downscaling

The term downscaling refers to the procedure of transferring large-scale information from GCMs to a regional or local scale, whereby the spatial resolution of the data is increased. Downscaling provides refined output at a higher spatial resolution, which is able to explicitly represent sub-grid scale heterogeneities. Consider precipitation as an example: because of the limited representation of regional topography and poor representation of mesoscale processes in GCMs, the spatial variations in precipitation intensity on regional scales often cannot be resolved. Therefore, downscaling plays a substantial role in mountainous areas, where precipitation patterns are strongly dependent on orography (Groppelli et al. [34]).

To bridge the gap between GCM output and the high-resolution climate variables required for hydrological modeling, various downscaling techniques have been developed. They can be classified into statistical downscaling (SD) and dynamical downscaling (DD). The basis for creating downscaled climate variables (e.g. temperature or precipitation) using SD is the underlying assumption that statistical relationships can be established between atmospheric processes occurring at different spatial scales [35]. These statistical relationships are used for downscaling, for instance, using weather typing schemes, transfer functions or weather generators [36]. SD is flexible and computationally cheap but is based on the assumption that the utilized statistical relationships do not change over time. Owing to the low computational cost of SD, many realizations are possible, which is for example useful when sampling uncertainties related to internal variability. DD, on the other hand, involves the use of higher-resolution (10–50 km) RCMs for limited regions (GCMs are run over the entire globe). These RCMs are run using boundary conditions provided by GCMs or reanalysis data. By using RCMs, DD offers a more physically realistic basis to downscaling when compared to SD because RCMs explicitly resolve mesoscale atmospheric processes that produce, for instance, heavy rainfall [37]. When a GCM is used to force an RCM (also called "nesting" and within this article referred to with the acronym GCM–RCM), regional detail is provided which is generally consistent with the driving GCM and also spatially coherent. For many parts of the world, climate change scenarios simulated by different RCMs are already freely available from public databases, e.g. through the CORDEX project for CMIP5 projections [38] or the ENSEMBLES project for CMIP3 projections [39]. There is no central archive for CORDEX, however, CORDEX data can be accessed from different portals, see: http://cordex.org/data-access/.

In some research papers, the terms SD and bias correction (bias correction is described in detail within Section 3.4) are used interchangeably or are instead referred to as simply "downscaling" or "statistical transformations" e.g. see the following papers which use these different terms in synonymous ways: Sunyer et al. [40], Fang et al. [41] or Gudmundsson et al. [42]. However, SD and bias correction have separate uses in some contexts, as evidenced by their motivation. The origin of SD and bias correction both date back to their use within numerical weather prediction (NWP). The first SD methods were implemented in the late 1940s [43], while bias correction developed some decades later. During the mid twentieth century NWP forecasts were too coarse to forecast weather variables at a local scale. SD models were, therefore, used to infer a statistical relationship between large-scale observational information (predictor) and an observed local variable (predictand). The statistical model was then applied to downscale the large-scale NWP forecast to the local-scale. In this circumstance, the large-scale NWP forecast is assumed to be perfectly fitted to large-scale observations. However, archived forecasts showed that the forecasts deviated heavily away from observations. Therefore, model output statistics (MOS) was introduced as a separate method from SD to correct for model biases. MOS infers a statistical relationship between a large-scale modeled predictor and an observed local-scale variable. Model biases already enter into the statistical relationship during MOS, allowing it to account for these biases. The terms MOS and bias correction are both found within current literature, where bias correction can be considered a subcategory of MOS. For further discussion of the origins of the terms bias correction, MOS, SD and their relatedness we refer the reader to Chapters 3 and 12 within the book by Maraun [44].

#### 3.3.1 Uncertainties Related to Downscaling

*Statistical downscaling*: Given that statistical relationships are established between observed and climate-modeled data, uncertainties related to observational datasets will influence the effectiveness of the SD technique. Section 5.2 discusses methods to accommodate for uncertainties stemming from observational datasets. In addition, the use of SD includes the uncertainty that results from the assumption that large-scale predictors are able to capture the climate change signal. This is discussed in greater detail within Section 6.3. Numerous SD methods exist and uncertainties can be introduced depending on the method used. It is, therefore, common to use multiple methods to accommodate for these uncertainties; see Section 4 for guidance on ensemble methods. For a review of different SD methods, see for instance Fowler et al. [45] or Maraun et al. [46].

*Dynamical downscaling*: Although their resolution is finer, RCMs are affected by issues that also affect GCMs and by the uncertainties listed in Section 3.2 (i.e. model structure/parameterization, natural variability). This causes systematic model errors, implying that there is the need to further post-process RCM data with bias correction and to use an ensemble approach (see Section 4) when using RCM simulations as input for modeling future streamflow [47, 48].

### 3.4 Bias Correction

In the context of this article, a bias is defined as "the systematic difference between a modeled property of the climate system and the corresponding real property" [49]. Such properties include mean temperature or summer precipitation. Bias correction is the process of correcting climate model output to reduce the effects of systematic errors in the climate models and to make the output more suitable as driving data for hydrological models. In the case where bias correction is applied between climate model output and observational data of different resolutions, then the bias correction also inherently closes the scale gap. Whether it is reasonable to use bias correction as a downscaling measure depends on the variable being downscaled, the difference in resolution, the study location, the bias correction method, and the statistical climate aspects that could be affected [44]. Initially, biases between observed and simulated climate variables over a historical period are identified (during the so-called "control run"). These biases then serve as a basis for establishing a transformation algorithm, which is used to correct both control and scenario driven RCM runs. This implicitly assumes that the biases are invariant over time, which is not always the case [48]. Although bias correction is usually performed on RCM data output, it should be noted that a direct bias correction of GCM data is also possible and likely computationally cheaper. However, the finer RCM resolution better resolves the regional-scale variability [50], which is beneficial especially in complex topography [51].

A large number of bias correction approaches have been developed to adjust climate model simulations as reviewed by Maraun et al. [50], Teutschbein and Seibert [52] or Chen et al. [53]. They can be classified according to their degree of complexity (i.e. how many statistical moments they are able to correct), ranging from simple scaling factors to more sophisticated methods such as quantile mapping. Among the different bias correction methods, quantile mapping (also referred to in the literature as distribution mapping, probability mapping, SD and histogram equalization) has been identified as the most efficient in adjusting RCM simulations. The idea behind this approach is to match the distribution of the RCM-simulated climate values with the observed distribution with the help of transfer functions (Figure 2). This has been shown to be superior to other bias correction methods because it is able to correct quantile dependent biases including wet day frequencies and intensities. The aforementioned bias correction methods can be considered "direct methods." The delta change approach is another widely used method to correct RCM data. The delta change method is considered separate from direct methods since it uses observations as a basis and then perturbs the observed time series rather than

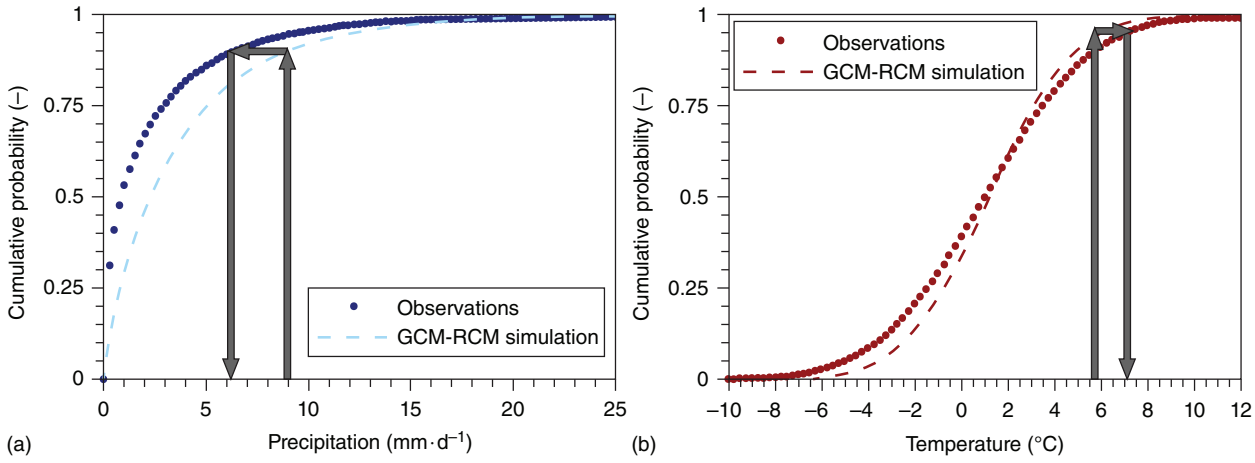



**Figure 2** Illustration of quantile mapping for (a) precipitation (gamma distribution) and (b) temperature (Gaussian distribution). In both cases, the distribution fitted to GCM-RCM-simulated values (dashed line) is transformed to fit the distribution fitted to the observed data (circles).

applying a correction to modeled data. The change factors are derived from changes simulated by the climate models. The statistical characteristics of the new time series remain the same as the observed time series, which makes it a stable and robust method. However, since the projected changes (i.e. the difference between simulated future and simulated past) are simply superimposed upon the observed data, it produces future time series with dynamics similar to current conditions. This implies that the delta change approach does not account for potential future changes in climate dynamics (e.g. the number of dry vs. wet days) and that major events (e.g. heavy precipitation or hot days) will change by the same amount as all other events (e.g. drizzle or cold days), which makes the delta change approach less suitable in hydrological climate change impact studies (Teutschbein and Seibert [48]). However, due to its simplistic nature this method is still very popular in the literature and can also easily be used within a classroom for teaching purposes (see Section 5.6).

Typically, bias correction methods are univariate, i.e. they adjust only one RCM-simulated variable at a time without guaranteeing consistency in spatiotemporal fields and different climate variables [50]. Bias correction has been shown to have a moderate effect on inter-variable relationships [54]. However, this remains an understudied aspect of hydrological climate change impact studies, and the use of a multivariate method versus a univariate method may lead to different conclusions in an impact study. For instance, when temperature and precipitation are jointly corrected within glacierized catchments, this may lead to more snowfall due to more precipitation falling while temperatures are below 0 °C as compared to a univariate approach [55]. Different multivariate methods are beginning to appear in the literature, for instance, Cannon [56] introduced a method based on the N-dimensional probability density function transform, which was originally used as an image processing technique. The technique combines univariate quantile mapping and random orthogonal rotations to match the multivariate distributions of climate model data to that of observed data. Another method called Multivariate Recursive Quantile Nesting Bias Correction (MRQNBC; [57]) corrects attributes of individual variables that result in a correction of the dependence biases between different variables. The Frequency Bias Correction (FBC; [58]) method is based on the concept that the variance of the time series can be expressed as a function of frequency. The biased time series is converted into the frequency domain using the forward Fourier transform and the peaks and phases are matched with that of the observational time series. In addition, a few attempts have been made to improve the physical links between bias-corrected variables by introducing copula-based correction methods [59–61]. Copulas are used to link univariate marginal distribution functions to form a multivariate distribution function. These two-dimensional methods are, however, not yet technically mature as they either do not establish a rigorous statistical relationship between the variables or are not able to correct data at a daily timescale.

#### 3.4.1 Uncertainties Related to Bias Correction

*Symptom vs. origin*: A major criticism of bias correction methods is that they only target the symptoms of model imperfections (i.e. biases in the simulations) and not the origins of these imperfections [49]. This leads to concerns about the ability of bias correction methods to correct future biases in a robust way. In a sense, bias correction provides the right answer (i.e. simulations looking like observations) but not necessarily for the right reasons. In addition, bias correction does not create subgrid variability [50] and assumes a stationarity of the bias (see Section 6.3 for further discussion). Despite these limitations, bias correction methods are still essential for hydrological impact studies, because without bias correction, systematic biases of raw climate model output would lead to substantial errors in hydrological projections.

*Bias correction method*: Studies have shown that the choice of bias correction method can also contribute to the total uncertainty of the modeling chain [40, 62]. For instance, Sunyer et al. [40] compared eight methods to downscale precipitation output (including four bias correction methods) from 15 RCMs from the ENSEMBLES project. Their results showed that the differences between the methods vary according to the catchments and the season being analyzed.

*Observational datasets*: Bias correction establishes statistical relationships between observed and modeled data. Therefore, uncertainty related to observed datasets will influence the effectiveness of the bias correction technique applied. Section 5.2 discusses methods related to the processing of observational dataset uncertainties.

### 3.5 Hydrological Models

Hydrological models are a simplification of real-world catchments and aim at representing the dominant hydrological processes. Hydrological models vary in their complexity ranging from purely empirical black box models to fully distributed physically based models [63]. For use in climate impact studies, bucket-type models, such as the HBV model [64, 65], are commonly used, as they are often considered to have sufficient complexity to capture the dominant hydrological processes, yet their data requirements are relatively modest. These models, also called conceptual models within the

hydrological community, are based on some physical reasoning and represent catchment processes by several interconnected buckets, which mimic water storage and transfer within a landscape. These models typically represent a catchment in a lumped or semi-distributed way. In lumped models, the catchment is considered to be spatially homogeneous, while in semi-distributed models, the heterogeneity within the catchment is accounted for using subunits. For snow-dominated catchments, for instance, the division into elevation zones is crucial, as it allows the model to account for changes in temperature and precipitation with elevation. The data requirements of these models include precipitation, temperature, and evapotranspiration which are usually sufficient for discharge modeling. Another advantage of bucket-type models is that they are simple enough to be easily applied and allow, due to their smaller computational demand, a more thorough uncertainty analysis. However, there are some instances where the use of a distributed process-based model is necessary to realistically capture the variable or process of interest. For a review of applications of process-based models in hydrology, we refer the reader to Fatichi et al. [66]. When starting with the production of streamflow projections, it is advisable to use a bucket-type hydrological model. Such models are easier to setup and run than process-based models, as they have lower data requirements (this is particularly important in data-scarce regions) and their run time is lower, enabling a larger ensemble of projections. For a thorough review of the different types of hydrological models and the pros and cons of different model complexities, we refer the reader to Hrachowitz and Clark [63]. For a historical overview of the development of catchment modeling see Todini [67].

#### 3.5.1 Uncertainties Related to Hydrological Models

*Parameters*: Parameter uncertainty is generally caused by (i) assumed stationarity of parameter values under changing climatic conditions and (ii) the difficulty to constrain model parameters using available data and knowledge. These two aspects of parameter uncertainty are limitations within hydrological climate change impact research and are further discussed in Sections 6.3 and 6.4, respectively.

*Model structure*: The response of the hydrological system to climate change can be impacted by model choice [63, 68, 69]. In practice, simpler models (i.e. lumped bucket-type models) often perform at least as well as the more complex models with regard to catchment discharge (Breuer et al. [70]), and more complexity does not guarantee that a model performs better under changed conditions [71–73]. However, this does not imply that models should not be improved; research has shown that improving process representation could increase model transferability into future conditions [74]. Yet, adding more complexity without the necessary data to support the additional scheme or parameter could lead to an increase in uncertainty along with slower run times. See Section 4.2 for a discussion of how to analyze model structure and Section 6.1 for limitations related to sampling within the hydrological model structure space.

*Observational datasets*: Hydrological models rely on observational data both as input to drive the simulations and for comparing the simulated time series of discharge for calibration. However, observational networks can contain uncertainty stemming from (i) instrument errors, (ii) errors in the conversion of relating measured values to the variable of interest (e.g. rating curve for discharge observations), (iii) spatial heterogeneities of the variable of interest (representativeness of the variations in the sample across space), and (iv) temporal variability of the variable of interest (whether sample variations are captured by temporal sampling) [75]. Since it is common practice to repetitively compare simulations to observation for calibration and validation purposes, issues within the observational network can lead to improper model setup and interpretation of results. Hydrological climate change impact studies rely on various observational data networks, which have different sources of uncertainty (see the earlier discussion) depending on the variable considered. For instance, although simulated discharge stemming from a calibrated hydrological model may match well with observed discharge, it is important to keep in mind that streamflow measurements during floods are uncertain. Please see Section 6.6 for a discussion of limitations related to observational dataset uncertainties.

## 4 Uncertainty Sampling and Decomposition

### 4.1 Ensemble Approach

It is now standard for studies investigating climate change impacts to rely on an ensemble approach [76]. This means that multiple runs rather than a single model run are performed and that the result is a range of possible outcomes rather than a single simulation. The various simulation runs differ in one or several aspects (e.g. different emission scenarios, climate models, climate model members, downscaling methods, bias correction methods, hydrological models, and parameter sets). Depending on computational resources and data availability, an ensemble can consist of a few (~10) to many (1000 or more) ensemble members. The different simulation runs are then often aggregated by computing the mean or median change and the spread

in the projections. The aggregation of the ensemble runs is usually a more robust estimate of future changes than if one would use one single run. Note that although both the mean and median can be used to combine ensemble runs, the median is often preferred over the mean since it is less affected by individual outliers, such as poorly performing models. The spread of the ensemble provides an estimate of the uncertainty in the projections. The following sections describe three ways to use ensembles to account for different forms of uncertainty: (i) uncertainties in the formulation of the models (structural uncertainty), (ii) uncertainties in parameter values within the models (parametric uncertainty) or (iii) uncertainty due to natural climate variability. Ensembles are also often used to isolate the contribution of other forms of uncertainty by utilizing different emission scenarios, bias correction methods, observational forcing, among other components of the modeling chain.

### 4.2 Ensembles to Test Structural Uncertainty

The assessment of model structural uncertainty is generally performed by using different model structures and characterizing the range of their output, known as the multi-model approach (Jiang et al. [77]; Hublart et al. [78]; Seiller et al. [79]). This method is applied in both climate and hydrological modeling. Hydrological models are developed by different modeling groups, hence their modeling philosophies and therefore structures vary (even amongst bucket-type models). Similarly, GCMs and RCMs are developed at different institutes, resulting in different representations of Earth system processes from one model to another (although often not entirely, see Section 6.2). This leads to a spread in the projections, representing model uncertainty. Figure 3 illustrates this concept for discharge projections: The spread of the projections was produced by forcing HBV using different GCM–RCMs.

### 4.3 Ensembles to Characterize Parametric Uncertainty

In a hydrological modeling context, the perturbed parameter approach consists of running the hydrological model multiple times using different parameter sets, generated using, for instance, Monte-Carlo procedures [80, 81], Bayesian methods [82, 83], evolutionary algorithms [84], or depth functions [85].

Within climate and hydrological modeling, the perturbed-parameter approach (also known as the perturbed physics ensembles [86]) involves the perturbation of model parameters (typically those poorly constrained by observational data) or parameterization schemes, thus creating separate simulations using each variant. This is done in order to test the model system sensitivity to the perturbations and to develop a range of equally likely model responses consistent with uncertain parameters/schemes. In the simplest form of this analysis, a single parameter is identified and the model is run. This parameter is then changed and the model is then rerun. The collection of the climate model runs as a whole is defined as an ensemble of different realizations. Typically, this approach entails the simultaneous modification of several parameters to evaluate their combined impact on the system and to estimate the range of uncertainty related to their prescribed values.

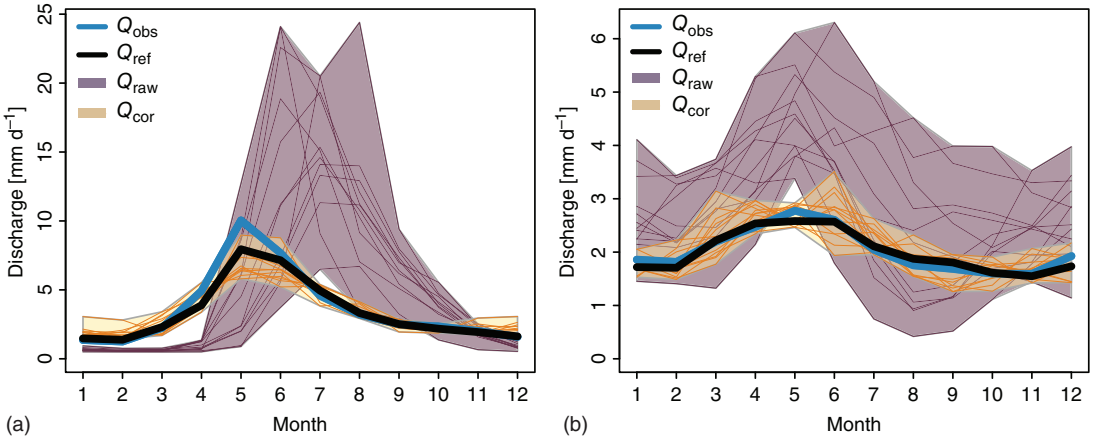



**Figure 3** Seasonal discharge for the Allenbach catchment (a) and Guerbe catchment (b), both located in Switzerland, for the time period of 1980–2010. Within each figure, observed discharge ($Q_{obs}$) is compared to simulated discharge driven by: (1) observed atmospheric forcing ($Q_{ref}$), (2) raw GCM-RCM data ($Q_{raw}$), and (3) bias-corrected GCM-RCM data ($Q_{cor}$). For both $Q_{raw}$ and $Q_{cor}$, the spread of the streamflow simulated using 12 GCM–RCMs from the coordinated regional climate downscaling experiment (CORDEX) is plotted (each line corresponds to one GCM-RCM, while the shading indicates the spread of the ensemble). Note how applying the bias correction reduces the spread among the members of the ensemble.

### 4.4 Ensemble Approach to Characterize Natural Variability

Recent studies showed that the importance of natural variability has often been underestimated when interpreting climate model projections. Initial condition ensembles involve the use of the same model and forcing but different start dates. Because of the chaotic nature of the climate system, small changes in temperature, humidity, etc. can result in highly different realizations of the system. This approach is, therefore, only applied to stochastic models or models with a stochastic setting. The chaotic nature of the atmosphere amplifies these slight differences, which results in some spread among the ensemble members. The spread provides a quantitative estimate of the natural variability of the climate system, often associated to noise.

Initial condition ensembles are utilized in a modeling setup like Deser et al. [33], who used one GCM, one emission scenario, identical initial conditions in the ocean, land and sea-ice components with different initial conditions in the atmospheric model. In this setup, the different model runs show the range of climate realities that can be achieved in the model world as a result of natural variability. Their study revealed considerable differences among the ensemble members, as a consequence of natural variability. In other words, even if models were perfect and future emissions known, the projections would still be uncertain, because of the chaotic nature of the atmosphere. A similar method was employed by Zhuan et al. [87], who used 29 different GCMs as well as a 40-member ensemble from one particular GCM. The differences among the 40 members are used to study the role of internal climate variability. The 40-member ensemble is further compared to the 29 GCM ensemble to estimate the timing of when human-induced climate change stands apart from internal climate variability. Other methodologies to analyze natural variability have also been employed, see Vidal et al. [88] and Fatichi et al. [89] who used a combination of a SD method and hydrological modeling. Natural variability represents a challenge to hydrological climate change impact studies because of its irreducible nature.

### 4.5 Techniques to Decompose Projection Uncertainty

The uncertainty in the projected hydrological changes can be decomposed and assigned to the different sources of uncertainty described previously. This requires a carefully developed experimental design, typically relying on the factorial combination of the different elements of the model chain (including all possible combinations of the model elements). It is a computationally demanding task, but it is necessary to isolate the contributions of each element of the model chain to the total uncertainty. Uncertainty partitioning is most commonly performed using an analysis of variance (ANOVA, Hawkins and Sutton [26]; Bosshard et al. [90]; Eisner et al. [91]). In an ANOVA framework (see Figure 4), the uncertainty is estimated from the variance among the ensemble members and the contribution of the elements of the model chain is additive. An uncertainty assessment of this kind allows for the determination of which elements of the model chain cause the most uncertainty, which helps with the design of future impact assessments. Including an additional hydrological model, for instance, might barely influence the projections, so additional runs can be avoided, which is helpful if computing resources are limited. Figure 4 shows an example of the outcome of an ANOVA decomposition where climate models are responsible for most of the spread in the projections, although there is some dependency on the variable, future period, and catchment of interest.

Studies such as Wilby [93] have suggested that the relative contribution of uncertainty from each step of the modeling chain to the final discharge projection is dependent on catchment characteristics. Addor et al. [92] produced streamflow projections for six Swiss catchments and showed that in nonglacierized catchments, uncertainty was mainly caused by GCM–RCMs. In contrast, in partially glacierized catchments, hydrological models played an equivalent role in discharge uncertainty. Bosshard et al. [90] showed that the time of year can also impact the contributions of uncertainty from different sources. They performed a variance decomposition on discharge projections and identified the GCM–RCMs to be the dominant source of uncertainty in the summer and autumn. Toward the end of the century, in winter and spring, the role of GCM–RCMs was found to diminish and instead hydrological models (as well as post-processing methods) become more important. Besides considering different catchment characteristics and the time of year, the contribution of uncertainty from the modeling chain is also dependent on which aspects of the hydrograph one evaluates. Meresa et al. [94] considered three sources of uncertainty: climate models obtained from EURO-CORDEX, hydrological model parameters achieved by calibration using observed streamflow over a reference period, and the process of fitting distribution models to extreme flow time series. The uncertainty of the hydrological parameters was estimated using the generalized likelihood uncertainty estimation (GLUE) approach. An ANOVA analysis showed that for low-flow extremes the uncertainty stemming from the hydrological parameters can be greater than the uncertainty from the climate models and distribution fitting process. For high-flow

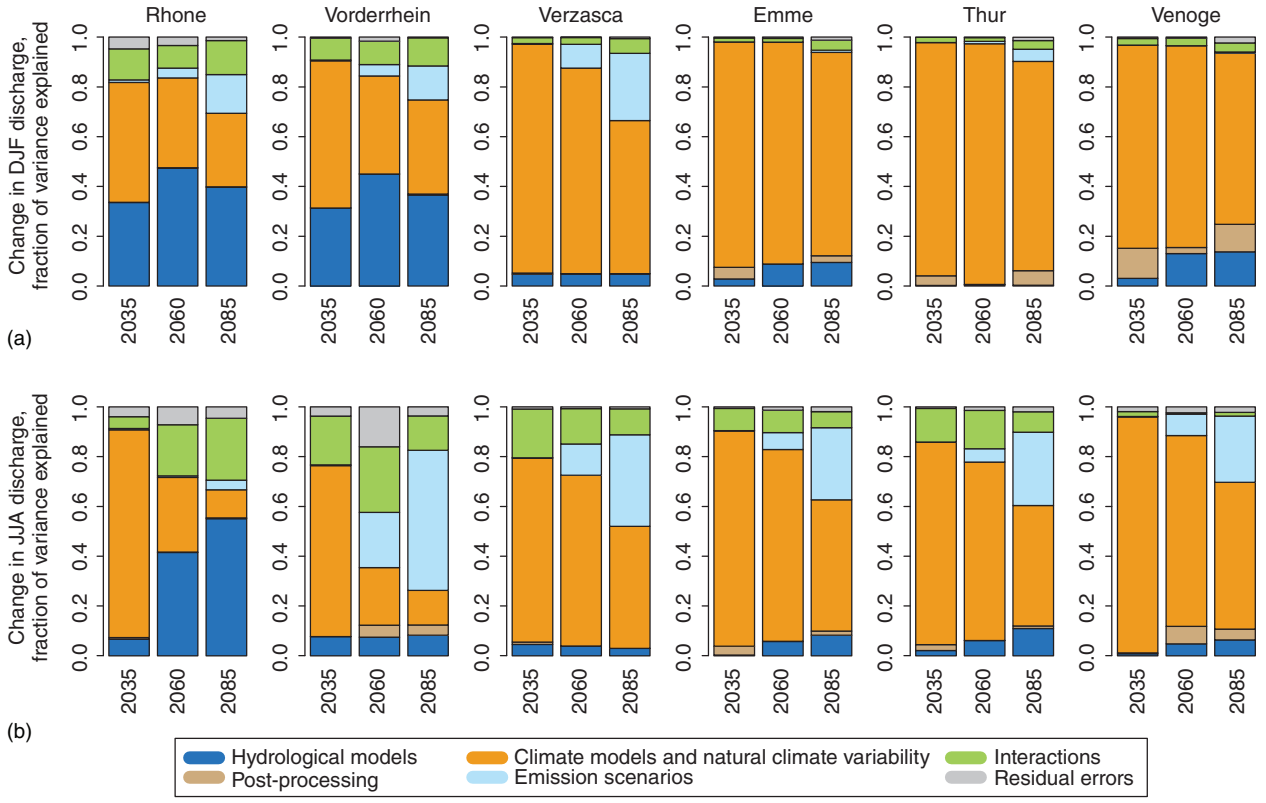



**Figure 4** Decomposition of the projection variance. ANOVA partitioning among the four selected sources of uncertainty, the significant interactions, and the residual errors. Results for discharge changes in (a) winter (DJF) and (b) summer (JJA) are shown for the three 30-year future periods, centered on 2035, 2060, and 2085 for six Swiss catchments. Source: Figure provided by Addor et al. (2014) [92].

extremes, they found the climate models to be the greatest contributor to total uncertainty.

## 5 Application of the Modeling Chain

### 5.1 Design of the Modeling Chain

Modeling chains can quickly become computationally demanding, and it is beyond a single modeling study to account for all uncertainties. System sensitivity tests (changing one or a few components of the model chain and leaving the other components unchanged) are often conducted (see Section 4 on uncertainty decomposition). A key decision is which emission scenario(s), climate model(s), bias correction technique(s), and hydrological model(s) to involve. The next subsections summarize the main steps involved in the production of hydrological projections.

### 5.2 Collection and Processing of Observed and Modeled Data

Besides the climate model simulations, the required data for a hydrological climate change impact study typically include observed precipitation, temperature, potential evaporation, and streamflow time series.

Meteorological observational data are often available as either station data or as a gridded product, which are derived from station data using interpolation techniques [95]. Incomplete or unavailable observational data are a common concern in climate change studies, which can be somewhat overcome by advanced methods to fill in the data [96] (e.g. interpolation, hindcasting). Since both precipitation and temperature vary with elevation, values always correspond to a certain reference elevation. For station data, this is simply the elevation of the station. For gridded datasets (including RCMs), however, values correspond to the mean elevation of the grid cell. If the study area is located in a region of complex topography, such as the Alps, temperature, and precipitation values derived from a gridded product should be corrected to account for the difference of the mean elevation of the grid cell and that of the actual elevation of the study area. This is especially important for gridded data with coarse resolution such as that of RCMs. Some hydrological models can automatically account for such differences by applying an additive correction to temperature and

a multiplicative correction to precipitation. Otherwise, this correction can also be performed by the modeler outside of the hydrological model. In addition, any systematic bias stemming from elevation differences is implicitly corrected through bias correction (see Section 3.4). However, it is important to note that when ranking the performance of raw RCM simulations, some of the biases within the data stem from elevation differences. When comparing raw RCM simulations, it is suggested to correct for elevation differences separately from bias correction, so that individual raw RCMs are not punished for issues related to elevation [97].

Potential evapotranspiration is not usually measured directly, but rather estimated using methods ranging from simple temperature-based equations such as Hargreaves and Blaney–Criddle (Xu and Singh [98]), to more physically based methods like the Penman–Monteith formulation [99]. Parsimonious formulations, such as temperature-based formulations, are useful in that they require less data compared to physically based formulations, but a downside of temperature equations is that they may be overly sensitive to climate change [100]. For a review on formulations used for estimating evaporation, McMahon et al. [101].

### 5.3 Estimation of the Hydrological Model Parameters

Parameters of a bucket-type model represent effective average values and most often cannot be linked to a specific property in the catchment that can be measured directly. Instead, calibration is used to estimate parameter values. Hydrological models are commonly calibrated against observed discharge, which means that a certain goodness-of-fit measure, such as Nash–Sutcliffe efficiency (NSE, Nash and Sutcliffe [102]) and Kling–Gupta efficiency (Gupta et al. [103]), is optimized by changing the parameter values. This can be done manually but usually automated methods are used (e.g. Monte Carlo simulations, genetic algorithms, Golberg [104]; Duan et al. [105]).

Once the model has been calibrated, it is important to test it over a period of time not used for the calibration, a step usually called validation. One particular challenge in climate impact studies is that the model is used to simulate streamflow under future climate, for which no observations are available. The assumption is often made that the same parameter values are still valid for the new situation. It is important to evaluate whether this assumption is reasonable and whether the model is "stable" under the respective change. This can be done using a differential split sample test (DSST). Using a DSST approach [106] for model evaluation is a suitable approach to explore the impacts of the assumed parameter stationarity on the simulations, i.e. to explore the transposability in time of the calibrated model. DSST relies on the calibration and validation of a hydrological model using sub-periods with contrasted climate conditions. The idea behind DSST is that the errors made by extrapolation from a time period with certain climate conditions to a time period with different conditions (both time periods using observational data so that the extrapolation errors can be quantified) can be used as a basis to determine whether the model will perform well under future climatic conditions [107]. Coron et al. [108], developed a generalized version of DSST (general split sample test: GSST), that allows for a large number of calibration-validation exercises by generating sub-periods systematically using a sliding window over the reference period. The main variables used to define the contrasted condition of sub-periods for DSST are precipitation (Vaze et al. [109]; Seiller et al. [110]; Tramblay et al. [111]; Ruelland et al. [112]), temperature (Hartmann and Bárdossy [113]), potential evapotranspiration (Coron et al. [108]), and discharge (Seibert [107]).

### 5.4 Evaluation and Bias Correction of the Climate Models

Until recently, subsets of GCM–RCMs have normally been chosen based on their ability to replicate current climate (temperature and precipitation metrics; Mendlik and Gobiet [114]; Wilcke and Bärring [115]). In this case, the evaluation of the GCM–RCMs and their bias correction takes place prior to hydrological modeling (Section 5.5). Recent studies such as Dalelane et al. [116] have targeted how representative the ensemble spread is by selection. Such a study is based on the notion that a climate ensemble is not fully robust because model results are not truly independent, as described in Section 6.2. The methodology first rejects poorly performing models and then selects the most independent models from the remaining ones. Besides evaluating a climate model's performance based on its ability to accurately represent climate variables, the use of hydrological metrics has recently gained more attention (e.g. high and low discharge) as evaluation criteria for the performance of GCM–RCMs and their bias correction. Streamflow can be considered an integrator of all atmospheric variables over a watershed. This approach allows for an instantaneous focus on atmospheric data with the largest influence on the simulated discharge [97]. With this approach, the evaluation of performance of GCM–RCMs and their bias correction is to be performed after hydrological modeling (Section 5.5) using hydrological metrics as standards for evaluation. It has been shown that climate model evaluation based on only

a single variable/metric is not sufficient [117] and that a robust selection should rather include the evaluation of a wide variety of variables/metrics.

The selection of an appropriate bias correction method is dependent upon multiple factors. The following list summarizes the most important considerations:

1) Can the model bias be considered time invariant (should a method which alters the climate change signal be applied)? A trend preserving method is justified if the bias can be considered time-invariant and conversely a nontrend preserving method can be applied if the method can be assumed to correct the time invariance of the bias.
2) Is downscaling to a higher resolution required? If so, the downscaling should capture local variations and the response to climate change.
3) How should the bulk of the climate distribution be corrected? Consider the correction of spatial, temporal, and multivariate aspects [50].

For a deeper discussion of this topic, Chapter 12, Section 10 of the book by Maraun [44] includes background information on the points listed above and provides a decision tree for the selection and evaluation of bias correction methods (Figure 12.17). The stepwise procedures to implement a particular bias correction are highly dependent on the method employed. Teutschbein and Seibert [52] provide an overview of the steps to carry out the following bias correction methods: linear scaling, local intensity scaling, power transformation, variance scaling, distribution mapping (i.e. quantile mapping), and the delta-change approach. The authors implemented these different methods and evaluated their ability to correct GCM-RCM temperature and precipitation for five Swedish catchments. The bias-corrected GCM-RCM data was then used to force a hydrological model to create discharge simulations for present and future conditions. Their results show that quantile mapping outperforms the other methods in that it corrects the most statistical characteristics and has the narrowest variability ranges.

### 5.5 Force Hydrological Model with Bias-corrected Climate Simulations and Analyze Streamflow

Once the hydrological model has been run with bias-corrected GCM-RCM data using parameters from Section 5.3, the discharge projections can then be analyzed. The general approach to analyze catchment discharge is as follows:

1) Apply the hydrological model for current conditions using observed historical precipitation and temperature ($PT_{obs}$) as forcing data. This usually implies some form of model calibration against observed discharge ($Q_{obs}$). This results in a time series of simulated discharge ($Q_{ref}$) based on observed climate data ($PT_{obs}$). This step should be conducted first, in order to establish the reliability of the hydrological model.
2) Compare the observed time series for precipitation and temperature ($PT_{obs}$) with the raw simulations from the climate model for current conditions ($PT_{raw}$), also called control run. Identify any biases.
3) Use the calibrated model to simulate the discharge for current conditions using the simulated time series for precipitation and temperature from the climate model, $Q_{raw,}$ $PT_{raw}$. This step should be completed prior to bias correction in order to test whether the discharge is sensitive to biases in the climate model data. If biases in the climate model data are apparent in the discharge, then proceed to step 4. If biases are not apparent in both steps 2 and 3, step 4 can be skipped and references to $Q_{cor}$ and $PT_{cor}$ can be neglected in the following steps.
4) In most cases, biases in the climate simulations will significantly impact the streamflow simulations, making the bias correction of the raw simulations precipitation and temperature data ($PT_{cor}$) necessary.
5) Use the calibrated hydrological model to simulate the discharge for current conditions using the bias-corrected simulated time series for precipitation and temperature from the climate model, $Q_{cor,}$ $PT_{cor}$.
6) Compare the various pairs of observed and/or simulated discharge to assess different sources of error (see Figure 5 for an example of these time series):

a) $Q_{obs}$ to $Q_{ref}$: biases associated with the hydrological model.
b) $Q_{ref}$ to $Q_{raw}$: biases related to the climate model.
c) $Q_{raw}$ to $Q_{cor}$: effect of the bias correction method.
d) $Q_{ref}$ to $Q_{cor}$: performance of the model chain after bias correction.

To consider uncertainties it is recommended to perform each step multiple times (i.e. use an ensemble method) whenever possible/suitable in the steps above. In step 1 for instance, due to parameter uncertainty, it is a good practice to allow for different parameterizations, i.e. to compile an ensemble of suitable parameter sets to be used in the further analyses. Ensemble means and medians, as well as spread measures, are suitable methods to show results and uncertainties (see Section 4).

### 5.6 Materials Available to Get Started

The Supplementary Materials for this article are available at the website: https://www.geo.uzh.ch/en/units/h2k/Services/Encyclopedia-Climate-Change.html. When

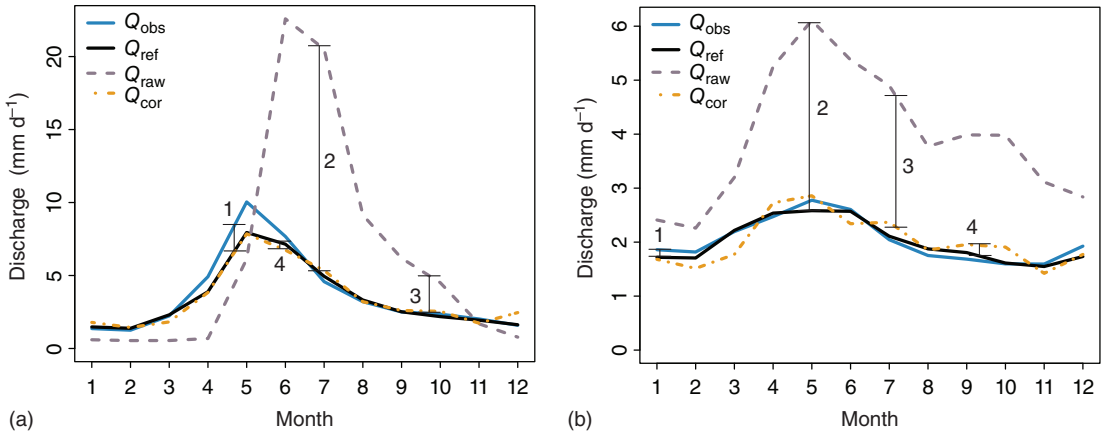



**Figure 5** Seasonal hydrographs for the Allenbach catchment (a) and Guerbe catchment (b), both located in Switzerland. Within each figure, observed discharge ($Q_{obs}$) is compared to simulated discharge driven by: observed climate data ($Q_{ref}$), raw RCM data ($Q_{raw}$), and bias corrected RCM data ($Q_{cor}$). The numbered lines refer to biases described in Section 5.5, step 6, where (1) $Q_{obs}$ to $Q_{ref}$: hydrological model biases, (2) $Q_{ref}$ to $Q_{raw}$: climate model biases, (3) $Q_{raw}$ to $Q_{cor}$: effect of the bias correction, and (4) $Q_{ref}$ to $Q_{cor}$: performance of model chain after bias correction.

beginning a hydrological climate change impact assessment, a necessary step is to download GCM-RCM data and to clip the data to the area of interest. This can be a challenge for those who are new to climate change projects. A guide is provided within the Supplementary Materials which walks the reader through the steps for pre-processing GCM-RCM data in NetCDF format.

For detailed instructions on how to apply a distribution-based bias correction method, the "qmap" package in R [42, 118] provides information on the application of quantile mapping. For those new to working with a hydrological model, the hydrological model HBV (Hydrologiska Byråns Vattenbalansavdelning; Bergström [64]; Bergstrom [119]; Seibert and Vis [120]) is available for download on the Supplementary Materials website. HBV is provided as a starting point given its successful implementation in classroom settings [120]. In addition, a simplified hydrological climate change impact project (using the delta change approach) is available in the Supplementary Materials. This project has been successfully used in a master's level course in the Department of Geography at the University of Zurich in the past years. The materials can be used as a guide for self-teaching or as a starting point for faculty wishing to assign a hydrological climate change research project in a classroom setting.

### 5.7 Potential Mistakes

It is important to note that GCM–RCM simulations cannot be used as weather conditions for "real" days. Climate models provide one possible realization of the climate evolution during a certain time period, and while this realization should reproduce the statistics of the observations, the individual values will be different from the observation. In other words, a GCM–RCM simulation for a certain day, say 1 January 1980 cannot and should not be compared to the observation on that specific date. Instead, GCM–RCM simulations should be evaluated in a climatological sense, by comparing long-term variables, such as the mean or variability over a long period (typically 20–30 years). Similarly, computing a goodness of fit measure such as NSE using daily discharge values does not make any sense when the hydrological model is driven by climate model output. An exception to this is when the RCM is driven by re-analysis data instead of a GCM.

Another potential mistake in the quantification of climate change impacts on discharge is the direct comparison of the discharge simulations driven by GCM–RCM output for the future with that of the observed discharge or the simulated discharge using observed atmospheric forcing. Both these comparisons are not suitable to quantify the effect of the simulated climate change because differences do not only arise from a changing climate but also include model errors. The appropriate approach is to compare the simulated discharge based on climate model outputs for both current and future conditions. By keeping the driving models the same between the historical and future time periods, the effects of climate change can be more easily isolated.

## 6 Limitations and Challenges

### 6.1 Sampling Within the Hydrological Model Structure Space

Most hydrological impact studies only involve a few different model structures that represent only a small sample of the multitude of model structures currently available. This poses a problem for two reasons. First, only a small part of the model structure space is sampled, meaning the spread in the projections may not reflect the full uncertainty in the future conditions. Second, these few models differ in so many respects that it is particularly difficult to determine which differences between them contribute most to the uncertainty in the projections. A methodology to diagnose differences in hydrological model structures is to use modular modeling frameworks, such as the Framework for Understanding Structural Errors (FUSE; [121]) or the structure for unifying multiple modeling alternatives (SUMMA; [18]).

### 6.2 Interdependence of Climate Model Structure

A main objective of combining the output from several climate models is to produce more robust projections. Some climate models perform better than others, leaving the researcher to decide whether to give higher weight to some models. The implementation of this can often be particularly difficult in that climate model groups often share code with one another and therefore the basis for the modeled physical processes can be very similar. As Knutti et al. [122] point out, new generations of GCMs often resemble their predecessors demonstrating that much of the physics and code remains the same through generations. In such cases, model agreement does not necessarily indicate correctness. The sharing of code means that models make similar assumptions of the physical system and therefore agreement amongst the models may come from an error shared by all models, and likewise, a low spread does not necessarily mean a low uncertainty. It could instead be the result of model interdependence, i.e. models relying on the same principles, sharing code, and being tuned using the same observations [30, 123]. In the current state of the research, there is no particular best practice approach to the combination of interdependent ensemble members.

### 6.3 Stationarity/Instability of Model Parameters

Both bias correction and SD methods are based on the assumption of stationarity, which implies that the correction algorithm or transfer function is assumed to also be valid under future conditions. In general, a transfer function/correction is derived based on the differences between the observational data and that of the climate model output over the historical period. For the future time period, the same transfer function/correction is applied. Maraun [124] analyzed the validity of the assumption of stationarity, by using an ensemble of GCM–RCMs to simulate present and future climate. All RCMs were forced by the same GCM, which was used to represent observed large-scale boundary conditions and a particular RCM was chosen to represent regional observations. By comparing the modeled simulations to the pseudo observational data, biases were found to be generally stable and bias correction was shown to considerably improve the future climate simulations. Yet, in some regions and for some seasons, bias correction was found to increase the future bias. While this will not be the case in all situations, bias correction usually reduces the biases of the raw RCM data even in the case of nonstationarity [54]. In addition, bias correction and SD are often applied only on a single temporal resolution (usually daily values); this does not ensure that multi-day statistics, which are essential for the modeling of droughts and high flows, are correctly captured [125].

Under nonstationary conditions, such as climate change, the effects of parameter uncertainty can be expected to be considerable (Coron et al. [108]; Poulin et al. [126]; Thompson et al. [127]). While parameter values in hydrological models theoretically should reflect the physical catchment characteristics and functioning, and not climatic conditions, several studies indicate that under climate change, parameter instability is mainly due to climate dependence of the calibrated parameter values [109, 128–130]. Caution is, therefore, needed when applying parameter values to different conditions. Several studies found that transferring parameters to different climate conditions resulted in significant uncertainties especially when moving to a drier and warmer climate [108, 109, 128].

### 6.4 Equifinality of Parameters

Parameter uncertainty is caused by the general difficulty of identifying a single "correct" parameter set. For instance, within bucket-type models, parameter values represent effective values at the catchment scale and are usually found by calibration (or regionalization of parameter values, which have been calibrated elsewhere) and many studies have demonstrated that it is impossible to identify one single "best" parameter set. This concept, also termed equifinality [131], means that there are multiple possible parameter sets, which perform similarly for a given calibration period but might result in significantly different results when being used for particular conditions, especially if these are outside the calibration

conditions [107, 132]. This provides a challenge to modelers because it implies that a certain level of understanding of hydrological processes cannot be gained through hydrological modeling as it exists today.

### 6.5 Climate Model Selection and Evaluation

At the time of this publication, there exists no general, all-purpose method to select and evaluate GCM–RCM combinations. Often selection is not in the control of the hydrological modeler since only some of the GCM–RCM combinations are available due to limited computing resources. A major difficulty in model selection and evaluation is the dependency on observational data as a benchmark (e.g. a model may perform well according to one observational dataset but poorly according to another). For instance, Gómez-Navarro et al. [133] and Kotlarski et al. [134] showed that the ranking of climate models differs depending on which observational datasets are used. Uncertainties in the observational dataset need to be smaller than the uncertainties stemming from the climate models so that climate models are not punished for the wrong reasons.

### 6.6 Accounting for Observational Uncertainties

A model, despite being firmly based on physical realism or empirically justified by performance, cannot produce accurate discharge predictions if forced with inaccurate data [135]. In many hydrological climate change impact studies, explicit consideration of observational dataset uncertainty is overlooked because it is often overshadowed by the more prominent biases and limitations associated with climate and hydrological models. It remains a major challenge for a practicing modeler to accommodate for imperfect observational data. In a practical sense, it is hardly possible to collect additional data for a particular model application. Interpolation, extrapolation, regionalization (i.e. relating information from a data-rich area to a data-poor area), and other more advanced methods are commonly used to accommodate for observational data deficiencies.

## 7 General Outlook

The main objectives of this article have been to introduce the topic of hydrological climate change impact modeling and to highlight how uncertainties are embedded within such research. The uncertainties discussed represent the areas within current research which have proven to be formidable roadblocks in the path to progress. Kundzewicz and Gerten [136] argue that uncertainty in projections of water resources have actually increased over time. This is due to the fact that with more information, we increasingly know more about what we do not know. Knutti and Sedláček [122] reiterate this idea by stating that uncertainties in climate change projections are not likely to be reduced quickly. Yet efforts to improve hydrological climate change impact modeling is of vital importance and progress should not be measured by how quickly model uncertainty decreases, but instead by how well we understand the processes driving climate change and its impacts.

Hydrological climate change research should be viewed with optimism. Although some uncertainties in projections may remain for the time being or cannot be reduced (e.g. natural variability), uncertainties should not prevent decisions from being made [122] nor deter those working on climate change impact research. Some of the most pressing research needs are those that may lead to more robust decisions and to a decrease in the uncertainty related to observations and projections of climate change [136]. Uncertainty can never be fully avoided and it is, thus, important to consider these uncertainties in decision making related to climate change impacts.

## Acknowledgments

The authors acknowledge the anonymous reviewer who provided helpful comments and thank in particular Sven Kotlarski who provided valuable insight into climate models and generously offered his feedback on this article.

## References

1 Alexander, L.V., Zhang, X., Peterson, T.C. et al. (2006). *J. Geophys. Res.* 111: 1–22. doi: 10.1029/2005JD006290.
2 Huntington, T.G. (2006). *J. Hydrol.* 319: 83–95. doi: 10.1016/j.jhydrol.2005.07.003.
3 IPCC Climate Change (2014). Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, IPCC, Geneva, Switzerland, Geneva, Switzerland. doi: 10.1017/CBO9781107415324.
4 Nijssen, B., O'Donnell, G.M., Hamlet, A.F., and Lettenmaier, D.P. (2001). *Clim. Chang.* 50: 143–175. doi: 10.1023/A:1010616428763.

5 Hirabayashi, Y., Mahendran, R., Koirala, S. et al. (2013). *Nat. Clim. Chang.* 3: 816–821. doi: 10.1038/nclimate1911.

6 Arora, V.K. and Boer, G.J. (2001). *J. Geophys. Res.* 106: 3335–3348. doi: 10.1029/2000JD900620.

7 Stahl, K., Hisdal, H., Hannaford, J. et al. (2010). *Hydrol. Earth Syst. Sci.* 14: 2367–2382. doi: 10.5194/hess-14-2367-2010.

8 Schuur, E.A.G., McGuire, A.D., Schädel, C. et al. (2015). *Nature* 520: 171–179. doi: 10.1038/nature14338.

9 Grover, V.I. *Managing Water Resources Under Climate Uncertainty*, (eds. S. Shrestha, A.P. Salam, A.K. Anal and M. van der Valk) 3–30. Cham/Heidelberg/New York/Dordrecht/London: Springer. doi: 10.1007/978-3-319-10467-6.

10 Várallyay, G. (2010). *Agronomy Research* 8, special issue II: 385–396. http://agronomy.emu.ee/category/volume-08-2010/special-issue-ii-volume-8-2010/.

11 Hall, J., Arheimer, B., Borga, M. et al. (2014). *Hydrol. Earth Syst. Sci.* 18: 2735–2772. doi: 10.5194/hess-18-2735-2014.

12 Djebou, D.C.S. and Singh, V.P. (2016). *Environ. Soc. Psychol.* 1: 36–49. doi: 10.18063/ESP.2016.01.002.

13 Wood, E., Roundy, J.K., Troy, T.J. et al. (2011). *Water Resour. Res.* 48: 1–10. doi: 10.1029/2010WR010090.

14 IPCC Climate Change (2013). The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge, UK and New York, NY, USA.

15 Teutschbein, C. and Seibert, J. (2010). *Geogr. Compass* 4: 834–860. doi: 10.1111/j.1749-8198.2010.00357.x.

16 Rojas, R., Feyen, L., Dosio, A., and Bavera, D. (2011). *Hydrol. Earth Syst. Sci.* 15: 2599–2620. doi: 10.5194/hess-15-2599-2011.

17 Shaad, K. (2018). *Hydrol. Sci. J.* 63: 1062–1077. doi: 10.1080/02626667.2018.1473871.

18 Clark, M.P., Nijssen, B., Lundquist, J.D. et al. (2015). *Water Resour. Res.* 51: 1–17. doi: 10.1002/2015WR017200.A.

19 Hagemann, S., Machenhauer, B., Jones, R. et al. (2004). *Clim. Dyn.* 2004: 547–567. doi: 10.1007/s00382-004-0444-7.

20 Dee, D., Fasullo, J., Shea, D., Walsh, J., and National Center for Atmospheric Research Staff. The Climate Data Guide: Atmospheric Reanalysis: Overview & Comparison Tables, https://climatedataguide.ucar.edu/climate-data/atmospheric-reanalysis-overview-comparison-tables (accessed 12 December 2016).

21 González-Zeas, D., Garrote, L., Iglesias, A., and Sordo-Ward, A. (2012). *Hydrol. Earth Syst. Sci.* 16: 1709–1723. doi: 10.5194/hess-16-1709-2012.

22 Fekete, B.M. and Vörösmarty, C.J. (2002). *Global Biogeochem. Cycl.* 16: 1042. doi: 10.1029/1999GB001254.

23 Brown, R.D., Brasnett, B., and Robinson, D. (2010). *Atmos. Ocean* 41: 1–14. doi: 10.3137/ao.410101.

24 Steger, C., Kotlarski, S., Jonas, T., and Schär, C. (2013). *Clim. Dyn.* 41: 735–754. doi: 10.1007/s00382-012-1545-3.

25 Räisänen, J. (2016). *Clim. Dyn.* 46: 339–353. doi: 10.1007/s00382-015-2587-0.

26 Hawkins, E. and Sutton, R. (2009). *Bull. Am. Meteorol. Soc.* 90: 1095–1107. doi: 10.1175/2009BAMS2607.1.

27 Climate change, IPCC (2001). *The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the IPCC*. Cambridge/New York, NY: Cambridge University Press.

28 McSweeney, R. and Hausfather, Z. (2018). *Carbon-Brief* 1–57.

29 Kour, R., Patel, N., and Krishna, A.P. (2016). *Arab. J. Geosci.* doi: 10.1007/s12517-016-2561-0.

30 Knutti, R., Masson, D., and Gettelman, A. (2013). *Geophys. Res. Lett.* 40: 1194–1199. doi: 10.1002/grl.50256.

31 Eyring, V., Bony, S., Meehl, G.A. et al. (2016). *Geosci. Model Dev.* 9: 1937–1958. doi: 10.5194/gmd-9-1937-2016.

32 Lorenz, E.N. (1963). *J. Atmos. Sci.* 20: 130–141. doi: 10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

33 Deser, C., Knutti, R., Solomon, S., and Phillips, A.S. (2012). *Nat. Clim. Chang.* 2: 775–779. doi: 10.1038/nclimate1562.

34 Groppelli, B., Bocchiola, D., and Rosso, R. (2011). *Water Resour. Res.* 47: 1–18. doi: 10.1029/2010WR009437.

35 Wilby, R.L. and Wigley, T.M.L. (1997). *Prog. Phys. Geogr.* 530–548. doi: 10.1177/030913339702100403.

36 Vaittinada Ayar, P., Vrac, M., Bastin, S. et al. (2016). *Clim. Dyn.* 46: 1301–1329. doi: 10.1007/s00382-015-2647-5.

37 Lafon, T., Dadson, S., Buys, G., and Prudhomme, C. (2013). *Int. J. Climatol.* 33: 1367–1381. doi: 10.1002/joc.3518.

38 Giorgi, F., Jones, C., and Asrar, G.R. (2009). *WMO Bull.* 58: 175–183.

39 Linden, P. van der, Mitchell, J.F.B. and Gilbert, P. (2009). ENSEMBLES: Climate Change and its Impacts: Summary of Research and Results from the ENSEMBLES Project, Met Office Hadley Centre, FitzRoy Road, Exeter EX1 3PB, UK. doi: PNR61, SIGE.

40 Sunyer, M.A., Hundecha, Y., Lawrence, D. et al. (2015). *Hydrol. Earth Syst. Sci.* 19: 1827–1847. doi: 10.5194/hess-19-1827-2015.

41 Fang, G.H., Yang, J., Chen, Y.N., and Zammit, C. (2015). *Hydrol. Earth Syst. Sci.* 19: 2547–2559. doi: 10.5194/hess-19-2547-2015.

42 Gudmundsson, L., Bremnes, J.B., Haugen, J.E., and Engen-Skaugen, T. (2012). *Hydrol. Earth Syst. Sci.* 16: 3383–3390. doi: 10.5194/hess-16-3383-2012.

43 Klein, W.H. (1948). *Bull. Am. Meteorol. Soc.* 439–453.

44 Maraun, D. and Widmann, M. *Statistical Downscaling and Bias Correction for Climate Research*. Cambridge: Cambridge University Press.

45 Fowler, H.J., Blenkinsop, S., and Tebaldi, C. (2007). *Int. J. Climatol.* 27: 1547–1578. doi: 10.1002/joc.1556.

46 Maraun, D., Wetterhall, F., Chandler, R.E. et al. (2010). *Rev. Geophys.* 48: 1–38. doi: 10.1029/2009RG000314.

47 Ivanov, M.A. and Kotlarski, S. (2016). *Int. J. Climatol.* doi: 10.1002/joc.4870.

48 Teutschbein, C. and Seibert, J. (2013). *Hydrol. Earth Syst. Sci.* 17: 5061–5077. doi: 10.5194/hess-17-5061-2013.

49 Maraun, D., Shepherd, T.G., Widmann, M. et al. (2017). *Nat. Clim. Chang.* 7: 764–773. doi: 10.1038/NCLIMATE3418.

50 Maraun, D. (2016). *Curr. Clim. Chang. Reports* 2: 211–220. doi: 10.1007/s40641-016-0050-x.

51 Giorgi, F., Torma, C., Coppola, E. et al. (2016). *Nat. Geosci.* 9: 584–589. doi: 10.1038/ngeo2761.

52 Teutschbein, C. and Seibert, J. (2012). *J. Hydrol.* 456–457: 12–29. doi: 10.1016/j.jhydrol.2012.05.052.

53 Chen, J., Brissette, F.P., Chaumont, D., and Braun, M. (2013). *Water Resour. Res.* 49: 4187–4205. doi: 10.1002/wrcr.20331.

54 Wilcke, R.A.I., Mendlik, T., and Gobiet, A. (2013). *Clim. Chan.* 120: 871–887. doi: 10.1007/s10584-013-0845-x.

55 Meyer, J., Kohn, I., Stahl, K. et al. (2018). *Hydrol. Earth Syst. Sci. Discuss.* 1–22. doi: 10.5194/hess-2018-317.

56 Cannon, A.J. (2018). *Clim. Dyn.* 50: 31–49. doi: 10.1007/s00382-017-3580-6.

57 Mehrotra, R. and Sharma, A. (2016). *J. Clim.* 3519–3539. doi: 10.1175/JCLI-D-15-0356.1.

58 Nguyen, H., Mehrotra, R., and Sharma, A. (2016). *J. Hydrol.* 538: 117–126. doi: 10.1016/j.jhydrol.2016.04.018.

59 Li, C., Sinha, E., Horton, D.E. et al. (2014). *J. Geophys. Res. Atmos.* 119: 153–162. doi: 10.1002/2014JD022514.

60 Piani, C. and Haerter, J.O. (2012). *Geophys. Res. Lett.* 39 (1–6): doi: 10.1029/2012GL053839.

61 Vrac, M. and Friederichs, P. (2015). *J. Clim.* 28: 218–237. doi: 10.1175/JCLI-D-14-00059.1.

62 Troin, M., Velázquez, J.A., Caya, D., and Brissette, F. (2015). *J. Hydrol.* 520: 268–288. doi: 10.1016/j.jhydrol.2014.11.047.

63 Hrachowitz, M. and Clark, M.P. (2017). *Hydrol. Earth Syst. Sci.* 3953–3973. doi: 10.5194/hess-21-3953-2017.

64 Bergström, S. *Development and Application of a Conceptual Runoff Model for Scandinavian Catchments*. Norrköping.

65 Lindström, G., Johansson, B., Persson, M. et al. (1997). *J. Hydrol.* 272–288. doi: 10.1016/S0022-1694(97)00041-3.

66 Fatichi, S., Vivoni, E.R., Ogden, F.L. et al. (2016). *J. Hydrol.* 537: 45–60. doi: 10.1016/j.jhydrol.2016.03.026.

67 Todini, E. (2011). *Hydrol. Res.* 42: 73. doi: 10.2166/nh.2011.096.

68 Seiller, G., Anctil, F., and Roy, R. (2017). *J. Hydrol.* 552: 313–340. doi: 10.1016/j.jhydrol.2017.07.002.

69 Ludwig, R., May, I., Turcotte, R. et al. (2009). *Adv. Geosci.* 21: 63–71. doi: 10.5194/adgeo-21-63-2009.

70 Breuer, L., Huisman, J.A., Willems, P. et al. (2009). *Adv. Water Resour.* 32: 129–146. doi: 10.1016/j.advwatres.2008.10.003.

71 Seibert, J. and (Ilja) van Meerveld, H.J. (2016). *Hydrol. Process.* 4966–4971. doi: 10.1002/hyp.10999.

72 Refsgaard, J.C. and Knudsen, J. (1996). *Water Resour. Res.* 32: 2189–2202. doi: 10.1029/96WR00896.

73 Coron, L., Andréassian, V., Perrin, C. et al. (2014). *Hydrol. Earth Syst. Sci.* 18: 727–746. doi: 10.5194/hess-18-727-2014.

74 Minville, M., Cartier, D., Guay, C. et al. (2014). *Water Resour. Res.* 50: 5044–5073. doi: 10.1002/2013WR013857.

75 Refsgaard, J.C., van der Keur, P., Nilsson, B. et al. (2006). *Hydrol. Earth Syst. Sci. Discuss.* 3: 1943–1985. doi: 10.5194/hessd-3-1943-2006.

76 Parker, W.S. (2013). *Wires Clim. Chang.* 4: 213–223. doi: 10.1002/wcc.220.

77 Jiang, T., Chen, Y.D., Xu, C. et al. (2007). *J. Hydrol.* 336: 316–333. doi: 10.1016/j.jhydrol.2007.01.010.

78 Hublart, P., Ruelland, D., Dezetter, A., and Jourde, H. (2015). *Hydrol. Earth Syst. Sci.* 19: 2295–2314. doi: 10.5194/hess-19-2295-2015.

79 Seiller, G., Roy, R., and Anctil, F. (2017). *J. Hydrol.* 547: 280–295. doi: 10.1016/j.jhydrol.2017.02.004.

80 Beven, K. and Binley, A. (1992). *Hydrol. Process.* 6: 279–298. doi: 10.1002/hyp.3360060305.

**81** Bastola, S., Murphy, C., and Sweeney, J. (2011). *Adv. Water Resour.* 34: 562–576. doi: 10.1016/j.advwatres.2011.01.008.

**82** Neuman, S.P. (2003). *Stoch. Environ. Res. Risk Assess.* 17: 291–305. doi: 10.1007/s00477-003-0151-7.

**83** Huard, D. and Mailhot, A. (2008). *Water Resour. Res.* 44: 1–19. doi: 10.1029/2007WR005949.

**84** Vrugt, J.A., Gupta, H.V., Bouten, W., and Sorooshian, S. (2003). *Water Resour. Res.* 39: doi: 10.1029/2002WR001642.

**85** Bárdossy, A. and Singh, S.K. (2008). *Hydrol. Earth Syst. Sci.* 12: 1273–1283.

**86** Bellprat, O., Kotlarski, S., Lthi, D., and Schr, C. (2012). *J. Geophys. Res. Atmos.* 117: 1–13. doi: 10.1029/2012JD018262.

**87** Zhuan, M.-J., Chen, J., Shen, M.-X. et al. (2018). *Hydrol. Res.* 49: nh2018059. doi: 10.2166/nh.2018.059.

**88** Vidal, J.-P., Hingray, B., Magand, C. et al. (2016). *Hydrol. Earth Syst. Sci.* 20: 3651–3672. doi: 10.5194/hess-20-3651-2016.

**89** Fatichi, S., Rimkus, S., Burlando, P., and Bordoy, R. (2014). *Sci. Total Environ.* 493: 1171–1182. doi: 10.1016/j.scitotenv.2013.12.014.

**90** Bosshard, T., Carambia, M., Goergen, K. et al. (2013). *Water Resour. Res.* 49: 1523–1536. doi: 10.1029/2011WR011533.

**91** Eisner, S., Flörke, M., Chamorro, A. et al. (2017). *Clim. Chan.* 141: 401–417. doi: 10.1007/s10584-016-1844-5.

**92** Addor, N., Rössler, O., Köplin, N. et al. (2014). *Water Resour. Res.* 50: 1–22. doi: 10.1002/2014WR015549.

**93** Wilby, R.L. (2005). *Hydrol. Process.* 19: 3201–3219. doi: 10.1002/hyp.5819.

**94** Meresa, H.K. and Romanowicz, R.J. (2017). *Hydrol. Earth Syst. Sci.* 21: 4245–4258. doi: 10.5194/hess-21-4245-2017.

**95** Bajamgnigni Gbambie, A.S., Poulin, A., Boucher, M.-A., and Arsenault, R. (2017). *J. Hydrometeorol.* 18: 247–264. doi: 10.1175/JHM-D-16-0032.1.

**96** Myers, D.E. (1994). *Geoderma* 62: 17–28. doi: 10.1016/0016-7061(94)90025-6.

**97** Hakala, K., Addor, N., and Seibert, J. (2018). *J. Hydrometeorol.* 19: 1321–1337. doi: 10.1175/JHM-D-17-0189.1.

**98** Xu, C.Y. and Singh, V.P. (2002). *Water Resour. Manag.* 16: 197–219. doi: 10.1023/A:1020282515975.

**99** Monteith, J.L. *Evaporation and Environment*. New York/Cambridge: Academic Press.

**100** Sheffield, J., Wood, E.F., and Roderick, M.L. (2012). *Nature* 491: 435–438. doi: 10.1038/nature11575.

**101** McMahon, T.A., Finlayson, B.L., and Peel, M.C. (2016). *WIRES Water* 3: 788–818. doi: 10.1002/wat2.1172.

**102** Nash, J.E. and Sutcliffe, J.V. (1970). *J. Hydrol.* 10: 282–290. doi: 10.1016/0022-1694(70)90255-6.

**103** Gupta, H.V., Kling, H., Yilmaz, K.K., and Martinez, G.F. (2009). *J. Hydrol.* 377: 80–91. doi: 10.1016/j.jhydrol.2009.08.003.

**104** Goldberg, D.E. (1989). *Genetic Algorithms in Search Optimization & Machine Learning*. Addison-Wesley Publishing Company, Inc. doi: 10.1007/3-540-44673-7.

**105** Duan, Q.Y., Gupta, V.K., and Sorooshian, S. (1993). *J. Opt. Theory Appl.* 76: 501–521. doi: 10.1007/BF00939380.

**106** Klemeš, V. (1986). *Hydrol. Sci. J.* 31: 13–24. doi: 10.1080/02626668609491024.

**107** Seibert, J. (2003). *Nord. Hydrol.* 34: 477–492.

**108** Coron, L., Andréassian, V., Perrin, C. et al. (2012). *Water Resour. Res.* 48: 1–17. doi: 10.1029/2011WR011721.

**109** Vaze, J., Davidson, A., Teng, J., and Podger, G. (2011). *Hydrol. Process.* 2612: 2597–2612. doi: 10.1002/hyp.8030.

**110** Seiller, G., Anctil, F., and Perrin, C. (2012). *Hydrol. Earth Syst. Sci.* 16: 1171–1189. doi: 10.5194/hess-16-1171-2012.

**111** Tramblay, Y., El Adlouni, S., and Servat, E. (2013). *Nat. Hazards Earth Syst. Sci.* 13: 3235–3248. doi: 10.5194/nhess-13-3235-2013.

**112** Ruelland, D., Hublart, P., and Tramblay, Y. (2015). *IAHS-AISH Proc. Rep.* 371: 75–81. doi: 10.5194/piahs-371-75-2015.

**113** Hartmann, G. and Bárdossy, A. (2005). *Adv. Geosci.* 5: 83–87.

**114** Mendlik, T. and Gobiet, A. (2016). *Clim. Change* 135: 381–393. doi: 10.1007/s10584-015-1582-0.

**115** Wilcke, R.A.I. and Bärring, L. (2016). *Environ. Model. Softw.* 78: 191–201. doi: 10.1016/j.envsoft.2016.01.002.

**116** Dalelane, C., Früh, B., Steger, C., and Walter, A. (2018). *J. Appl. Meteorol. Climatol.* 57: 477–491. doi: 10.1175/JAMC-D-17-0141.1.

**117** Jury, M.W., Prein, A.F., Truhetz, H., and Gobiet, A. (2015). *Am. Meteorol. Soc.* 28: 5575–5582. doi: 10.1175/JCLI-D-14-00430.1.

**118** Gudmundsson, A.L. (2012). *Package 'qmap'*. doi: 10.5194/hess-16-3383-2012.bernexp.

**119** Bergstrom, S. (1992). The HBV Model – Its Structure and Applications, SMHI Reports RH No. 4, Norrköping.

**120** Seibert, J. and Vis, M.J.P. (2012). *Hydrol. Earth Syst. Sci.* 16: 3315–3325. doi: 10.5194/hess-16-3315-2012.

**121** Clark, M.P., Slater, A.G., Rupp, D.E. et al. (2008). *Water Resour. Res.* 44: 1–14. doi: 10.1029/2007WR006735.

**122** Knutti, R. and Sedláček, J. (2013). *Nat. Clim. Chang.* 3: 369–373. doi: 10.1038/nclimate1716.

**123** Tebaldi, C. and Knutti, R. (2007). *Philos. Trans. A. Math. Phys. Eng. Sci.* 365: 2053–2075. doi: 10.1098/rsta.2007.2076.

**124** Maraun, D. (2012). *Geophys. Res. Lett.* 39: 1–5. doi: 10.1029/2012GL051210.

**125** Addor, N. and Seibert, J. (2014). *Hydrol. Process.* 28: 4823–4828. doi: 10.1002/hyp.10238.

**126** Poulin, A., Brissette, F., Leconte, R. et al. (2011). *J. Hydrol.* 409: 626–636. doi: 10.1016/j.jhydrol.2011.08.057.

**127** Thompson, J.R., Green, A.J., and Kingston, D.G. (2014). *J. Hydrol.* 510: 259–279. doi: 10.1016/j.jhydrol.2013.12.010.

**128** Dakhlaoui, H., Ruelland, D., Tramblay, Y., and Bargaoui, Z. (2017). *J. Hydrol.* 550: 201–217. doi: 10.1016/j.jhydrol.2017.04.032.

**129** Merz, R., Parajka, J., and Blöschl, G. (2011). *Water Resour. Res.* 47: 1–17. doi: 10.1029/2010WR009505.

**130** Ruelland, D., Guinot, V. and Levavasseur, F. (2009). *Proceedings of Symposium HS.2 at the Joint IAHS & IAH Convention*.

**131** Beven, K. (2006). *Hydrol. Process.* 20: 3141–3146. doi: 10.1002/hyp.6396.

**132** Seibert, J. (1997). *Nord. Hydrol.* 28: 247–262. doi: 10.2166/nh.1997.015.

**133** Gómez-Navarro, J.J., Montvez, J.P., Jerez, S. et al. (2012). *Geophys. Res. Lett.* 39: 1–5. doi: 10.1029/2012GL054206.

**134** Kotlarski, S., Szabó, P., Herrera, S. et al. (2017). *Int. J. Climatol.* doi: 10.1002/joc.5249.

**135** McMillan, H., Jackson, B., Clark, M. et al. (2011). *J. Hydrol.* 400: 83–94. doi: 10.1016/j.jhydrol.2011.01.026.

**136** Kundzewicz, Z.W. and Gerten, D. (2015). *J. Hydrol. Eng.* 20: doi: 10.1061/(ASCE)HE.1943-5584.0001012.

PAPER II

# Hydrological Modeling to Evaluate Climate Model Simulations and Their Bias Correction

KIRSTI HAKALA

*Department of Geography, University of Zurich, Zurich, Switzerland*

NANS ADDOR

*Department of Geography, University of Zurich, Zurich, Switzerland, and Research Applications Laboratory, National Center for Atmospheric Research, Boulder, Colorado, and Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, United Kingdom*

JAN SEIBERT

*Department of Geography, University of Zurich, Zurich, Switzerland, and Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, Uppsala, Sweden*

(Manuscript received 3 October 2017, in final form 23 May 2018)

ABSTRACT

Variables simulated by climate models are usually evaluated independently. Yet, climate change impacts often stem from the combined effect of these variables, making the evaluation of intervariable relationships essential. These relationships can be evaluated in a statistical framework (e.g., using correlation coefficients), but this does not test whether complex processes driven by nonlinear relationships are correctly represented. To overcome this limitation, we propose to evaluate climate model simulations in a more process-oriented framework using hydrological modeling. Our modeling chain consists of 12 regional climate models (RCMs) from the Coordinated Downscaling Experiment–European Domain (EURO-CORDEX) forced by five general circulation models (GCMs), eight Swiss catchments, 10 optimized parameter sets for the hydrological model Hydrologiska Byråns Vattenbalansavdelning (HBV), and one bias correction method [quantile mapping (QM)]. We used seven discharge metrics to explore the representation of different hydrological processes under current climate. Specific combinations of biases in GCM–RCM simulations can lead to significant biases in simulated discharge (e.g., excessive precipitation in the winter months combined with a cold temperature bias). Other biases, such as exaggerated snow accumulation, do not necessarily impact temperature over the historical period to the point where discharge is affected. Our results confirm the importance of bias correction; when all catchments, GCM–RCMs, and discharge metrics were considered, QM improved discharge simulations in the vast majority of all cases. Additionally, we present a ranking of climate models according to their hydrological performance. Ranking GCM–RCMs is most meaningful prior to bias correction since QM reduces differences between GCM–RCM-driven hydrological simulations. Overall, this work introduces a multivariate assessment method of GCM–RCMs, which enables a more process-oriented evaluation of their simulations.

## 1. Introduction

Some of the most significant effects of climate change are expected to impact hydrological processes, such as snowmelt and timing of discharge (Salathé et al. 2007; Pechlivanidis et al. 2015). Therefore, it is of growing importance to create accurate projections of streamflow while understanding and reducing biases in the climate model projections. For the task of simulating streamflow at the catchment scale, it is common to employ a chain of models beginning with general circulation models (GCMs), which can then be statistically or dynamically downscaled, the latter by using regional climate models [RCMs; see Fowler et al. (2007) for a review of downscaling techniques]. Yet, even the latest generation of GCM–RCMs feature substantial biases (Terzago et al. 2017). Since streamflow is sensitive to changes in temperature and precipitation, even small biases can influence a

*Corresponding author*: Kirsti Hakala, kirsti.hakala@geo.uzh.ch

DOI: 10.1175/JHM-D-17-0189.1

system to the point of changing its normal dynamics (e.g., Li et al. 2014). GCM–RCM output is therefore usually bias-corrected prior to its use as input to a hydrological model (Themeßl et al. 2011b; Teutschbein and Seibert 2012; Räisänen and Räty 2013).

Streamflow is controlled by a wide range of hydrometeorological processes. When streamflow is simulated, the realism of the simulations reflects how well those processes are represented in models. Here we use hydrological modeling to evaluate the atmospheric forcing provided by a recent suite of GCM–RCM combinations.

For streamflow to be correctly simulated, the combination of the hydrologically important aspects of precipitation and temperature (including intervariable relationships) should be correct. However, compensating biases such as overly high summer temperature and precipitation amounts may still lead to realistic streamflow if evaporation is unrealistically large. Atmospheric variables should then be checked to make sure that their individual values are also realistic. The impact of bias correction on meteorological intervariable relationships has been previously studied. Wilcke et al. (2013) evaluated whether bias correction degrades or improves intervariable relationships between temperature, precipitation, relative humidity, wind speed, global radiation, and surface air pressure, using metrics such as autocorrelation and intervariable correlation. Their study comprised over 80 stations within Austria as well as 18 stations within Switzerland, and quantile mapping (QM) was used as a bias correction technique. QM removes quantile-dependent biases by transforming a climate simulation time series so that its cumulative distribution function corresponds to that of the observations (Gudmundsson et al. 2012; Maraun 2013). Wilcke et al. (2013) conclude that QM results in either improvement or has no clear effect on autocorrelation and no discernible effect on correlation between variables. This suggests that QM does not degrade intervariable dependencies. Li et al. (2014) investigated intervariable relationships using a bias correction method that explicitly accounts for the correlation between variables. After the application of their joint bias correction method, their results showed not only a reduction of biases in the mean and variance but also an improvement in the correlation between temperature and precipitation. Both Wilcke et al. (2013) and Li et al. (2014) use correlation to characterize the strength of the linear relationship between variables. However, many hydrological processes are not linear. Snowmelt, for instance, is rather a threshold-dependent process, and the accuracy of the simulations around 0°C is particularly important. Similarly, the interaction of antecedent wet conditions, rainfall intensity, and resulting discharge also exhibits threshold behaviors (Zehe and Sivapalan 2009).

To overcome the limitations associated with standard statistical evaluation tools, here we propose a more process-based investigation of climate model simulations using a modeling framework that captures the interactions between temperature and precipitation leading to discharge. We use this framework to rank climate models and to assess the influence of QM on the simulated discharge. Since streamflow inherently incorporates the dynamics between temperature and precipitation at the catchment scale, the evaluation of simulated discharge, with and without bias correction, can be used to determine if the relationship between meteorological variables is properly represented by climate models and how it is impacted by quantile mapping.

We use this evaluation framework to rank GCM–RCMs in order to support their selection for impact studies. Although it is essential to carefully select appropriate climatological data as input to hydrological models, choosing which GCM–RCM combinations to carry forward in the modeling chain is not always straightforward (Mendlik and Gobiet 2016). In practice, subsets of GCM–RCMs are generally selected based on their ability to replicate current climate, typically using temperature and precipitation metrics (e.g., Johnson and Sharma 2015). In addition to culling poorly performing models, model selection reduces the computational burden. As Wilcke and Bärring (2016) point out, full ensembles of GCM–RCM simulations can be too big for impact modelers to handle, and often specific GCM–RCMs are hand-picked. Mendlik and Gobiet (2016) argue that model performance under current climate should be used to remove extremely unrealistic models but not to make a selection of ''best performing'' models because it is unclear whether those specific models will provide the most realistic future projections. Although metrics to evaluate climate models have been established for some time, there is a lack of a standard index or procedure. Gleckler et al. (2008) used a wide set of metrics to evaluate 22 atmospheric variables simulated by 22 GCMs, focusing on global scales of the simulated mean annual cycle. They observed that the ranking of models varies considerably from one variable to the next, which points to the importance of considering a wide range of variables to comprehensively evaluate GCM performance. More recently, Jury et al. (2015) used a model performance index, developed by Reichler and Kim (2008), to evaluate the skill of GCMs according to their ability to reproduce near-surface and atmospheric variables. The index combines the climate model's performance at simulating multiple variables (e.g., surface and upper-air variables for temperature


FIG. 1. Map showing the locations of the eight Swiss study catchments in yellow and the underlying topography in gray. The hillshade topography is derived from a 25-m digital terrain model provided by the Swiss Federal Office of Topography (Swisstopo).

and precipitation). Their results show that there is little correlation between the performances of different variables, and thus their study also suggests that ranking GCMs based on a singular variable is inadequate. Here we use a wide variety of hydrological metrics to evaluate GCM–RCM combinations.

There are two main goals for this study. The first is to perform an evaluation of GCM–RCM simulations under current climate based on an integrated assessment of precipitation and temperature time series with respect to their hydrological significance. The methods used within this paper can be applied to evaluate climate models and their bias correction, regardless of the climate model or bias correction used. The ability (or inability) to correctly simulate streamflow is a way to assess the realism of the climate simulations. The second goal for this study is to rank GCM–RCMs based on how well they enabled us to capture hydrological variables. This research aims to provide modelers and end users with a new perspective on GCM–RCM performance that accounts for interactions between atmospheric variables (precipitation and temperature) at the catchment scale.

## 2. Data and methods

### *a. Study catchments and observational data*

Eight mesoscale catchments with areas ranging from 28 to 117 $km^2$ were selected as study catchments. They cover a wide range of regime types and elevations (Fig. 1, Table 1), with negligible human influences. The study catchments were also selected to have little to no glacial cover. Karstic topography is negligible in the majority of the catchments with the exception of the Breggia catchment, whose geology primarily includes permeable rock with sedimentary fissures. The Cassarate catchment was therefore selected as an additional study area for its similarities to the Breggia catchment and its lack of karstic topography. Research catchments in Switzerland are designated and managed by the Swiss Federal Office for the Environment (FOEN). Daily discharge data (24-h mean) were provided by the FOEN.

Meteorological data were retrieved from the gridded TabsD and RhiresD MeteoSwiss datasets. TabsD (Frei 2014) and RhiresD (Frei and Schär 1998; Schwarb 2000) are gridded daily temperature and precipitation data covering the domain of Switzerland. These gridded data products are available at a 2-km resolution and are based on daily temperature (mean of 10-min interval measurements) and precipitation totals measured (automatic and manual) at the high-resolution gauging network of MeteoSwiss, known as SwissMetNet (MeteoSwiss 2010). Note that the effective resolution of RhiresD is roughly 15–20 km or larger (approximate average interstation distance; MeteoSwiss 2013a). In regards to the TabsD data, there are particularly large errors in inner Alpine valleys (MeteoSwiss 2013b). Because of the lack of interpolation accuracy of the TabsD data in these areas, these cold air pool environments are systematically overestimated in winter. The interpolation errors are small for the other seasons.

TABLE 1. Main characteristics of the eight Swiss catchments including catchment area, karst percentage, elevation, glacier coverage, regime type, lapse rate, and precipitation gradient (calculated using MeteoSwiss data).

| Gauging station (ID) | Area ($km^2$) | Mean elevation (m MSL) | Glacier coverage (%) | Karst areas (%) | Lapse rate [°C (100 m)$^{-1}$] | Precipitation gradient [% (100 m)$^{-1}$] | Regime type |
|---|---|---|---|---|---|---|---|
| Murg–Wängi (2126) | 78.9 | 650 | 0 | 0 | −0.39 | 10.2 | Low elevation, rain influenced |
| Mentue–Yvonand (2369) | 105 | 679 | 0 | 0 | −0.33 | 9 | Jura Mountains, rain influenced |
| Guerbe Belp (2159) | 117 | 837 | 0 | 5 | −0.37 | 4.1 | High elevation, rain influenced |
| Breggia–Chiasso (2349) | 47.4 | 927 | 0 | 95 | −0.21 | 1.9 | High elevation, south facing |
| Cassarate–Pregassona (2321) | 73.9 | 990 | 0 | 0 | −0.43 | 1.9 | Rain/snow influenced, south facing |
| Sitter–Appenzell (2112) | 74.2 | 1252 | 0.08 | 0 | −0.36 | 4.4 | Transitional area between rain and snow |
| Allenbach–Adelboden (2232) | 28.8 | 1856 | 0 | 8 | −0.45 | 3.6 | Snow influenced, alpine catchment |
| Dischmabach–Davos (2327) | 43.3 | 2372 | 2.1 | 0 | −0.45 | 0 | Snow influenced with some glacierization |

### b. *GCM–RCMs*

#### 1) EURO-CORDEX

Daily temperature and precipitation series simulated by 12 RCMs, driven by five different GCMs (Table 2), were obtained from the Coordinated Regional Downscaling Experiment (CORDEX; www.cordex.org) via the CH2018 archive (http://www.ch2018.ch/en/home-2/). CORDEX is part of a collaborative modeling effort where GCM projections from the Coupled Model Intercomparison Project (CMIP5; https://cmip.llnl.gov/cmip5/) were downscaled using RCMs operated by different research institutes. Given our focus on Swiss catchments, GCM–RCMs were selected from the European domain of the CORDEX project (EURO-CORDEX; http://www.euro-cordex.net/). The list of the GCM–RCMs used in this study is provided in Table 2. For additional information regarding EURO-CORDEX climate modeling, we refer to Kotlarski et al. (2014), which provides an evaluation of ERA-Interim–driven EURO-CORDEX scenarios for Europe.

EURO-CORDEX provides simulations at 0.11° (~12.5 km) and 0.44° (~50 km) on a rotated grid. Given that the alpine domain was considered, only the higher-resolution 0.11° simulations were used within this study. The area of any given study catchment is smaller than the area of one RCM grid cell. Based on the orientation of a particular catchment and its relation to the RCM gridded system, typically 3–4 RCM grid cells contribute with some areal fraction to each catchment.

TABLE 2. Overview of the 12 EURO-CORDEX simulations used in this study. All models were run on a ~12.5-km grid. Bold text indicates the abbreviations used throughout the text and figures when referring to the models. The institutes of the models are indicated in standard font. The ensemble member information from the driving GCM is indicated by italics and parentheses, where "r" refers to the realization, "i" to the initialization method, and "p" to the physics version used.

| No. | GCM (member) | RCM | Calendar |
|---|---|---|---|
| 1 | CNRM-CERFACS-**CNRM-CM5** (*r1i1p1*) | CLMcom-**CCLM4-8-17** | Gregorian |
| 2 | ICHEC-**EC-EARTH** (*r12i1p1*) | CLMcom-**CCLM4-8-17** | Gregorian |
| 3 | MOHC-**HadGEM2-ES** (*r1i1p1*) | CLMcom-**CCLM4-8-17** | 360 |
| 4 | MPI-M-**MPI-ESM-LR** (*r1i1p1*) | CLMcom-**CCLM4-8-17** | Gregorian |
| 5 | CNRM-CERFACS-**CNRM-CM5** (*r1i1p1*) | SMHI-**RCA4** | Gregorian |
| 6 | ICHEC-**EC-EARTH** (*r12i1p1*) | SMHI-**RCA4** | Gregorian |
| 7 | IPSL-IPSL-**CM5A-MR** (*r1i1p1*) | SMHI-**RCA4** | No leap |
| 8 | MOHC-**HadGEM2-ES** (*r1i1p1*) | SMHI-**RCA4** | 360 |
| 9 | MPI-M-**MPI-ESM-LR** (*r1i1p1*) | SMHI-**RCA4** | Gregorian |
| 10 | ICHEC-**EC-EARTH** (*r1i1p1*) | KNMI-**RACMO22E** | Gregorian |
| 11 | ICHEC-**EC-EARTH** (*r3i1p1*) | DMI-**HIRHAM5** | Gregorian |
| 12 | IPSL-IPSL-**CM5A-MR** (*r1i1p1*) | IPSL-INERIS-**WRF331F** | Gregorian |

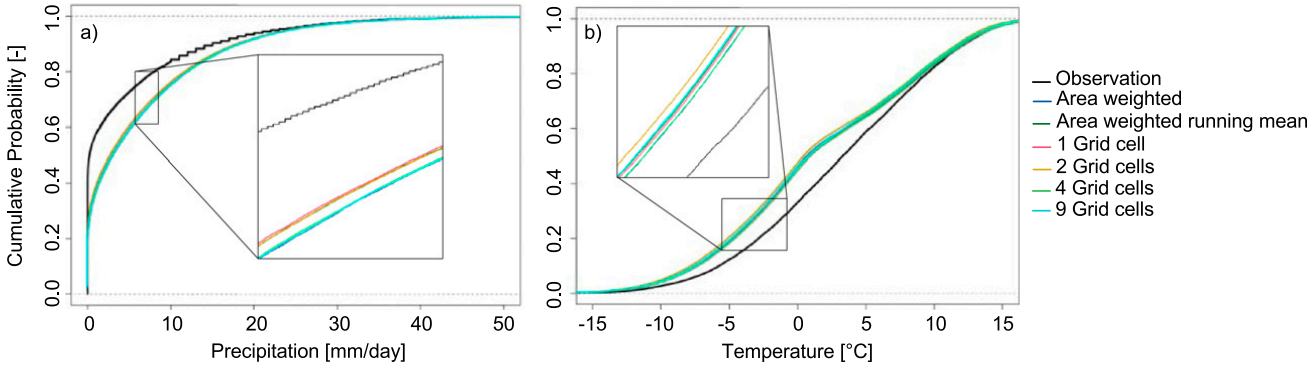



FIG. 2. CDFs for (a) temperature and (b) precipitation from one GCM–RCM (CNRM-CM5–CCLM4-8-17) for the Allenbach catchment.

One GCM–RCM (CM5A-MR–RCA4) uses a non-leap-year calendar (Table 2). For temperature and precipitation simulations from this GCM–RCM, the days before and after 29 February during leap years were used to interpolate the time series to a Gregorian calendar. The HadGEM2-ES–CCLM4-8-17 and the HadGEM2-ES–RCA4 models use a 360-day calendar; in this case the data were kept at a 360-day calendar, and the hydrological model was run using this calendar.

### 2) DATA EXTRACTION

Temperature, precipitation, and streamflow data were extracted for the catchments listed in Table 1 for the time period from 31 December 1979 to 31 December 2009 (from 30 December 1979 to 30 December 2009 in the case of the 360-day calendar GCM–RCMs). For each catchment, observational data were extracted using an area-weighted method, which comprised the following steps:

1) Identify all grid cells that overly the catchment.
2) According to the percent of overlap, a particular grid cell will be given a relative weight.
3) The precipitation and temperature time series are then extracted from the overlying grid cells, and the relative weight is applied to each grid cell's time series.
4) The average of all time series is then calculated, resulting in the area-weighted mean time series for the catchment.

A visual analysis was carried out to inspect different extraction methods of the GCM–RCM data, which involved extracting the 1) the closest grid cell to the centroid of the catchment, 2) the mean of the two closest grid cells, 3) the mean of the four closest, 4) the mean of the nine closest, 5) the area-weighted mean, and 6) the area-weighted running mean where each gridcell value is replaced by a $3 \times 3$ mean of the surrounding grid cells. Temperature and precipitation from all six extraction methods were compared to observational data using mean monthly averages, cumulative distribution functions (CDFs), and extreme high and low quantiles. Overall, the six methods delivered similar results (see example in Fig. 2). Therefore, an area-weighted mean was used to derive catchment mean values for both the gridded observational and GCM–RCM products.

### 3) BIAS CORRECTION

Bias correction techniques have been shown to be effective within different settings, such as the correction of daily GCM–RCM precipitation (Themeßl et al. 2011a), the improvement of simulated streamflow characteristics (Teutschbein and Seibert 2012), and enabling improved performance for the projection of temperature for the far future time period (Räisänen and Räty 2013). These studies and others also indicate that the QM method outperforms other simpler methods, such as the delta-change approach, local intensity scaling, and power transformation. In addition, nonparametric QM has been shown to have a higher skill in reducing biases in GCM–RCM precipitation compared to distribution-derived and parametric transformations (Gudmundsson et al. 2012).

For the purpose of this study, we do not explicitly differentiate between the biases of the GCM and RCM. Rather, the aggregated total bias (RCM biases and remnant biases from the GCM) was corrected by employing a nonparametric quantile transformation of seasonal distributions. Following a nonparametric method, CDFs were constructed for the following seasons using daily data: December–February (DJF), March–May (MAM), June–August (JJA), and September–November (SON) for both the observed and the GCM–RCM-simulated climate variables. The "qmap" package in R (Gudmundsson et al. 2012; Gudmundsson 2016) was used to map the CDF of the simulations onto the CDF of the observations.

### 4) SNOW ACCUMULATION IN EURO-CORDEX GCM–RCMS

Snow water equivalent (SWE) in EURO-CORDEX GCM–RCMs contains large biases compared to observational datasets. Terzago et al. (2017) analyzed

EURO-CORDEX SWE over the Greater Alpine Region and reported that several GCM–RCMs tend to constantly accumulate snow cover at high elevations. Therefore, for our study, snow depth was plotted for all GCM–RCMs and all catchments as well as for an additional 5–6 grid cells surrounding the catchments. Considering the area within the catchments and the surrounding grid cells, we found that snow towers are present in the EC-EARTH–RACMO22E simulations for the following two catchments: Dischmabach and Allenbach. These snow towers begin accumulating snow at the onset of the GCM–RCM simulation and reach an unrealistic height of more than 400 m by the end of the century (hereafter referred to as snow towers). Other GCM–RCMs may be affected by snow towers; however, such towers did not occur within or near our study catchments. Although snow is not explicitly provided as input to the hydrological model, the presence of snow towers may impact the temperature within the catchment and its change signal. Terzago et al. (2017) chose to eliminate all GCM–RCMs with unrealistic snow accumulation trends for use in future scenario analysis. For the purposes of this study, it was decided to evaluate all GCM–RCMs despite snow accumulation issues to test whether the snow towers have noticeable effects on catchment temperature and consequently discharge.

### c. *Hydrological modeling*

#### 1) HBV MODEL

The bucket-type Hydrologiska Byråns Vattenbalansavdelning (HBV) model (Bergström 1976; Lindström et al. 1997) was used to simulate daily streamflow values for each catchment. Here we used the version HBV-light (Seibert and Vis 2012). The HBV model relies on four routines: snow, soil, response, and routing routines. The HBV model is considered a semidistributed model since it allows for the catchment to be subcompartmentalized into different elevation zones, derived from a digital elevation model (DEM). As input, HBV requires temperature, precipitation, and potential evaporation. Within HBV, the flow of water through a catchment is represented in the following way: precipitation is first ingested as input, and HBV then simulates it as either rain or snow according to a threshold temperature within the ''snow routine.'' Next, the soil routine is activated where rainfall and snowmelt are divided into either the soil box or groundwater recharge depending on the water content of the soil box. Actual evaporation from the soil box equals potential evaporation when water availability is not limiting evaporation, and a linear reduction is used when water availability is limiting. Following the soil routine, the ''response function'' is activated where groundwater recharge is added to the upper groundwater box and percolates at a specific rate (defined by a model parameter) to a lower groundwater box. Runoff is then simulated as the sum of three linear outflows from the two boxes. Finally, within the ''routing routine,'' a triangular weighting function is applied to the generated runoff to represent the transport along the stream network. For additional model descriptions, we refer the reader to previous publications about the HBV model (Bergström 1976; Lindström et al. 1997; Seibert and Vis 2012). For the remainder of the text, the term HBV refers to the version HBV-light being used in this study.

#### 2) CALIBRATION AND VALIDATION OF HBV

The Lindström measure (Lindström et al. 1997), which is a combination of the model efficiency [Nash–Sutcliffe efficiency (NSE); Nash and Sutcliffe 1970] and volume error, was used as an objective function to calibrate HBV. The Lindström measure is computed as NSE minus 0.1 multiplied by the relative volume error and can range between $-\infty$ and 1. A value of 1 refers to a perfect match between modeled discharge and observed discharge. HBV was calibrated using a genetic algorithm and Powell optimization (GAP; Seibert 2000) method (5000 model runs for the genetic algorithm and an additional 1000 runs for the Powell optimization). The GAP optimization method works by selecting and recombining high-performing parameter sets with each other. At the conclusion of these runs, the parameter set associated with the highest objective value was selected. This process was repeated 10 times to produce 10 optimized parameter sets. Calibration was performed by first splitting the daily time series into two subsets. The first subset, 1980–94, was used to calibrate with a warmup period of one year, 1979. Validation was then performed on the second subset, 1995–2009, with a warmup period of one year, 1994. For the calibration period, model efficiency and Lindström measure values were above 0.7, and for the validation period, values above 0.6 were achieved for all catchments.

#### 3) CORRECTION FOR ELEVATION DIFFERENCE WITHIN HBV

To account for the difference between the elevation of the RCM grid cell(s) and that of the station observational network, we computed for each catchment the long-term mean monthly values of the temperature lapse rate and precipitation gradient using MeteoSwiss gridded data. All catchments show an annual cycle for both the temperature lapse rate and precipitation gradient (Fig. 3). Given that each catchment's observed values show significant deviations from the HBV default

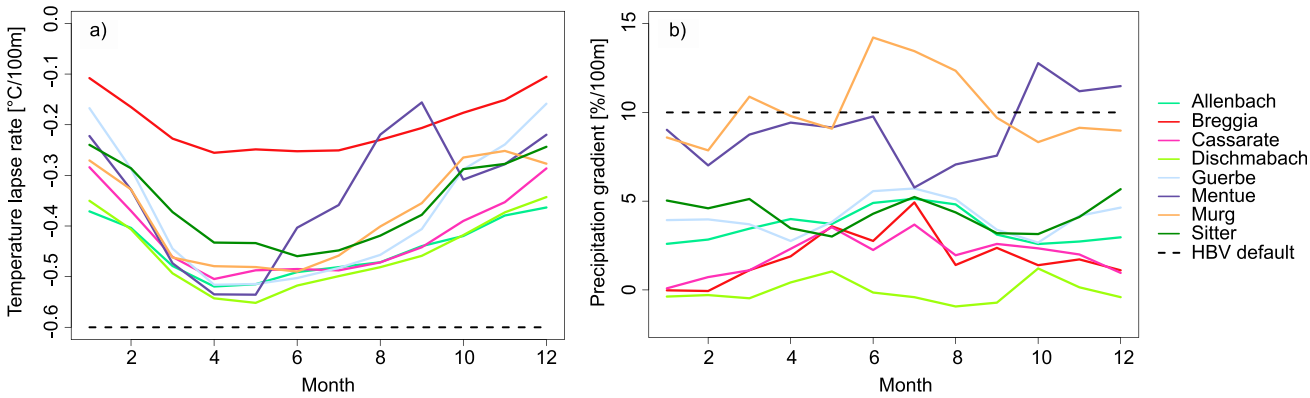



FIG. 3. Long-term monthly lapse rates for (a) temperature and (b) precipitation for all study catchments, which were used for the HBV simulations to reflect the topography of each catchment using elevation bands.

[temperature lapse rate of $-0.6°C\ (100\,m)^{-1}$ and precipitation gradient of 10% $(100\,m)^{-1}$], catchment-specific long-term mean monthly averages were used.

The temperature and precipitation catchment averages derived from the climate model simulations were adjusted (temperature was adjusted additively and precipitation was adjusted using a multiplicative relationship) to account for the difference in elevation between the RCM grid cells and the catchment elevation using these monthly constants. The climate variables were then bias corrected. Bias correction could have been used to correct for climate model biases without first correcting for elevation differences. By correcting for the elevation difference separately, the benefit of the bias correction can be isolated and the quality of uncorrected GCM–RCM simulations can be assessed without penalization because of the elevation of their grids.

#### 4) VALIDATION OF PERFORMANCE OF RCMS ACCORDING TO HYDROLOGICAL METRICS

The final step in the modeling chain is to run HBV using raw and bias-corrected GCM–RCM data as forcing and using the 10 parameter sets described in section 2c(2). In total, the streamflow series comprise:

- $Q_{obs}$, observed discharge monitored by FOEN;
- $Q_{ref}$, discharge simulated by HBV using MeteoSwiss forcing;
- $Q_{raw}$, discharge simulated by HBV using raw GCM–RCM data as forcing; and
- $Q_{qm}$, discharge simulated by HBV using QM GCM–RCM data as forcing.

The differences between $Q_{obs}$ and $Q_{ref}$ reflect errors in the atmospheric forcing and in HBV structure and parameter values. Differences between $Q_{ref}$ and $Q_{raw}$ reflect errors resulting from GCM–RCM biases. Differences between $Q_{raw}$ and $Q_{qm}$ reflect the impacts of the bias correction.

For all catchments and all climate models, the difference between parameter sets was smaller than the difference between $Q_{raw}$ and $Q_{ref}$. This indicates that the hydrological simulations are more sensitive to the bias correction than to the difference between the parameter sets. After quantile mapping, the difference between parameter sets becomes more important, as indicated by the observation that $Q_{qm}$ fits more closely to $Q_{ref}$ (Fig. 4). In the remainder of the paper, the simulations from the 10 parameter sets were averaged to produce a single discharge time series.

The following metrics were used to evaluate the simulations: long-term mean monthly discharge for the cold season (DJF) and warm season (JJA), low flow (Q5) and high flow (Q95), 7-day low flow, annual maximum, and the half-flow date (the day of the year when half the annual discharge has been measured). Given that some GCM–RCMs operate on different calendars (e.g., HadGEM2-ES–driven RCM models operate on a 360-day calendar), the half-flow date was calculated according to the number of days within the calendar's year. After a half-flow date was calculated for each individual year, the median of those values was then used. To alleviate any biased effects from extreme years, seasonal hydrological metrics (DJF and JJA) were each calculated by finding the mean value for each individual year and then further taking the median over all years. All other metrics (Q5, Q95, 7-day low flow, annual maximum) involved finding the annual value per year and then taking the median of all years. To standardize the various metrics, the relative error was calculated by comparing $Q_{ref}$ to $Q_{obs}$, $Q_{raw}$, and $Q_{qm}$:

$$E_{obs} = (Q_{obs} - Q_{ref})/Q_{ref},$$
$$E_{raw} = (Q_{raw} - Q_{ref})/Q_{ref}, \quad \text{and}$$
$$E_{qm} = (Q_{qm} - Q_{ref})/Q_{ref}.$$

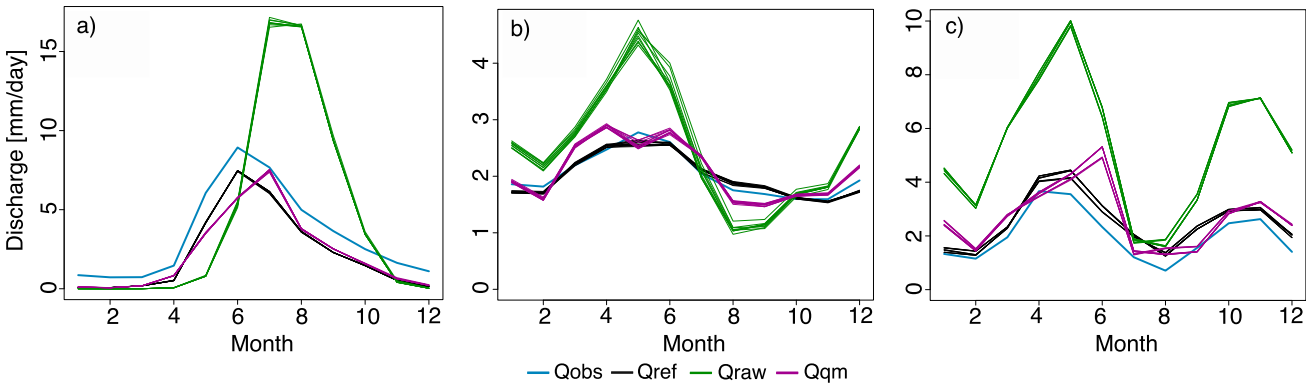



FIG. 4. Discharge for the time period of 1980–2009 for (a) the Dischmabach catchment and the GCM–RCM (CNRM-CM5–CLM4-8-17), (b) the Guerbe catchment and GCM–RCM (CNRM-CM5–RCA4), and (c) the Breggia catchment and GCM–RCM (HadGEM2-ES–RCA4). Ten simulations, which stem from the 10 parameter sets, are shown for each of the following: $Q_{ref}$, $Q_{raw}$, and $Q_{qm}$.

For the half-flow date metric, the absolute difference was used (by calculating $Q_{raw} - Q_{ref}$ and $Q_{qm} - Q_{ref}$ and $Q_{obs} - Q_{ref}$). The benefit(s) of QM can then be analyzed by comparing hydrological metrics of $E_{raw}$ versus $E_{qm}$. In addition, both raw and quantile mapped GCM–RCMs can then be ranked according to the performance of runoff simulations, which are based on the precipitation and temperature time series extracted from the GCM–RCMs. NSE is a commonly used metric to evaluate the realism of streamflow simulations. However, NSE is known to emphasize errors in large flows (Schaefli et al. 2007; Criss and Winston 2008). Large flows are only one part of the hydrograph and are not necessarily the main interest for all end users. Therefore, we considered various parts of the hydrograph that are likely to correspond to an end-user's interests.

### d. *Experimental design*

Overall, we combined 12 GCM–RCMs, 8 catchments, one hydrological model run with 10 parameter sets, and one bias correction method (both raw and bias-corrected data are used) and evaluate them over 1970–2009. In a factorial way, we analyzed 1920 discharge simulations [12 GCM–RCMs × 2 postprocessesing methods (raw and QM) × 8 catchments × 10 parameter sets]. In addition, we also analyzed 80 discharge simulations (8 catchments × 10 parameter sets) driven by observational forcing and 8 observational discharge datasets, leading to a total of 2008 discharge time series (Fig. 5).

## 3. Results

### a. *Evaluating individual effects of quantile mapping*

The first objective of this study was to explore whether QM reduces biases in hydrological simulations and how QM changes meteorological intervariable relationships. In particular, we investigated whether the amplitude and timing of the annual precipitation and temperature cycles are correctly captured and how this influences the annual discharge cycle simulated by HBV.

Prior to bias correction, biases in raw GCM–RCM precipitation $P_{raw}$ were substantial. In our study catchments, precipitation biases take the form of either a wet bias that persists primarily throughout the year (Fig. 6a) or a wet bias in the winter and spring months, often with a dry bias in the summer months (Figs. 6b,c). These main types of precipitation biases can also be seen within the other catchments not shown here. Bias-corrected precipitation $P_{qm}$ shows an improvement over $P_{raw}$ and generally fits more closely to observed precipitation $P_{obs}$. However, biases can still remain even after bias correction (see section 4a for further discussion). Additionally, temperature biases were present. Prior to QM, the largest biases in temperature $T_{raw}$ were found in high-elevation catchments. Within these catchments, a cold bias is evident for the entire annual cycle. Lower-elevation catchments are less severely affected. After bias correction, temperature $T_{qm}$ matches well with observed temperature $T_{obs}$ irrespective of the elevation of the catchment. The effect of these biases and bias correction on discharge depends on the elevation of the catchment.

In high-elevation catchments such as Allenbach and Dischmabach, the combination of wet biases in the winter/spring and a general cold bias often leads to a delay in discharge ($Q_{raw}$), peaking 1–1.5 months after both $Q_{obs}$ and $Q_{ref}$ (Fig. 6g). The delay in discharge is due to precipitation falling as snow in the winter months at these elevations. In addition, the magnitude of discharge is often much greater than both $Q_{obs}$ and $Q_{ref}$, which indicates an overestimation of snow accumulation. After QM, the cold biases are most often improved (Fig. 6d), and the wet precipitation bias in the winter and spring months is generally reduced (Fig. 6a). Therefore, both the timing and the magnitude of the resulting

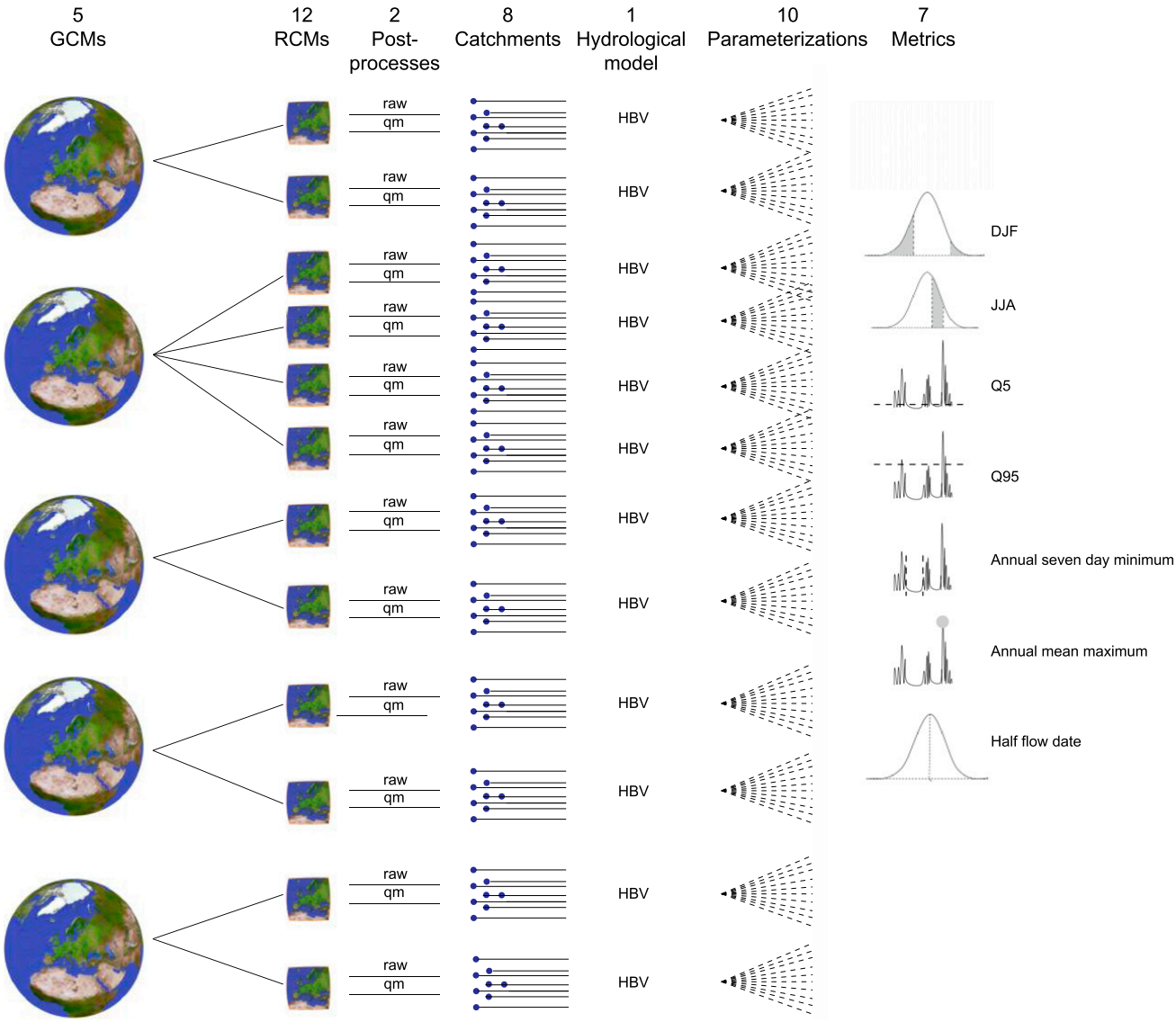



FIG. 5. Modeling chain of the general workflow used in this study.

discharge ($Q_{qm}$) matches more closely with both $Q_{obs}$ and $Q_{ref}$ (Fig. 6g).

In low- to medium-elevation catchments (679–1252 m MSL) such as Breggia, Sitter, and Murg, cold biases were less pronounced (Figs. 6e,f), but wet biases were often found (Figs. 6b,c). Depending on the catchment, these biases either persisted throughout the year or only impacted the winter and spring months, often with a dry bias in the summer months. This results in a bias in the magnitude of the discharge; however, timing in these mid- to low elevations is less affected compared to the high-elevation catchments (Figs. 6h,i).

A comparison of $E_{raw}$ to $E_{qm}$ is plotted side by side (noted as "variable" with raw or qm) in Figs. 7a–c. In the majority of cases, QM leads to a decrease of bias in hydrological variables: a striped pattern is visible when QM and raw results are displayed side by side (e.g., Q95 columns; Figs. 7a–c). This striped pattern indicates relatively high versus low percent error when comparing raw to QM simulations. Overall, QM increases the agreement of $P$ and $T$ time series with observations, which leads to an improvement of the runoff time series simulated by HBV. However, there are instances where QM does not result in an improved hydrologic performance (i.e., instances where the striped pattern is not present; see Fig. 7a, DJF columns). An explanation is that the relative percent error is very sensitive in low-flow metrics where small discharge values are compared to one another. Occasionally, such small differences can result in large relative errors. However, when compared to errors over the rest of the year, the errors over winter are rather small in absolute terms.

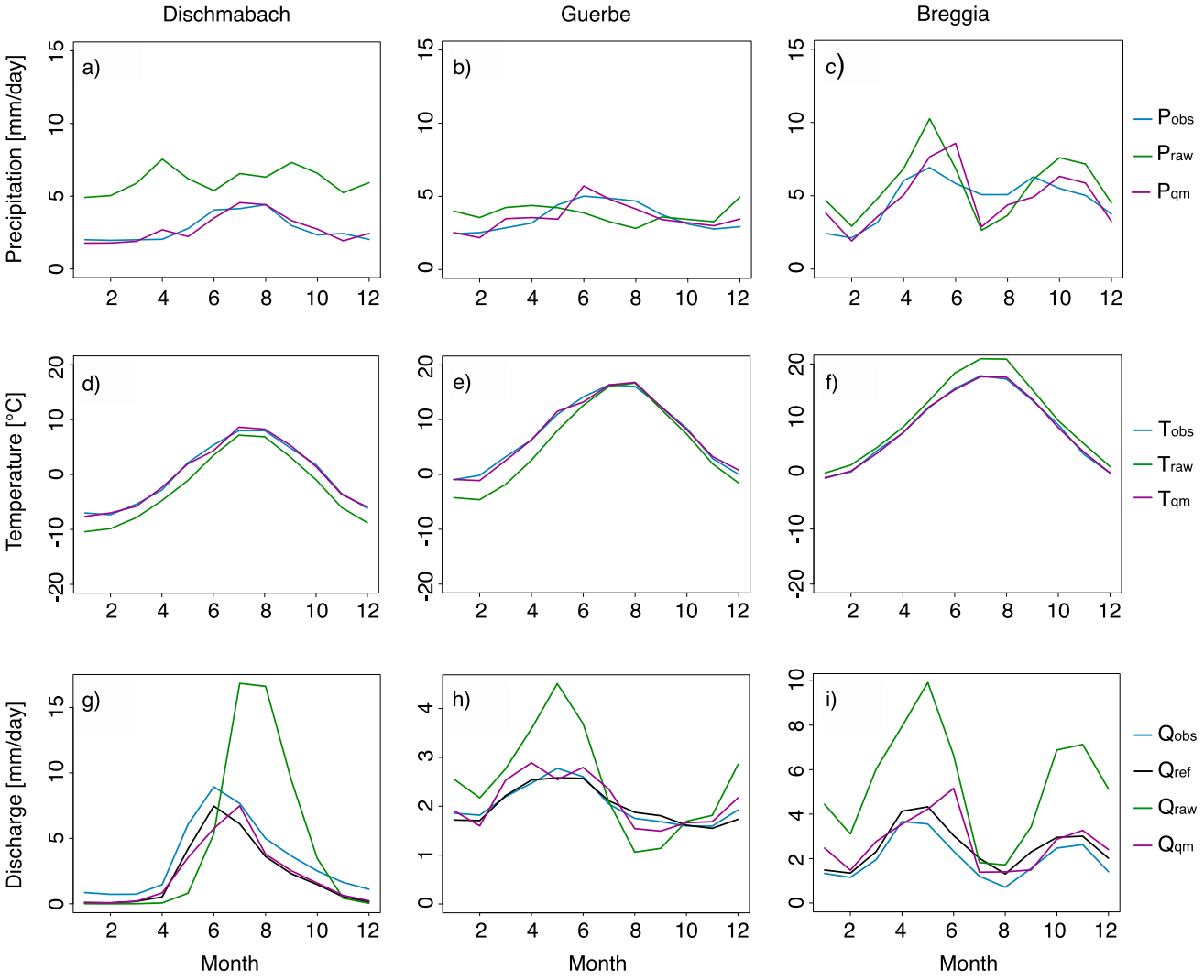



FIG. 6. The long-term mean monthly (a)–(c) precipitation, (d)–(f) temperature, and (g)–(i) discharge for three example catchments. The data from one GCM–RCM are used for each catchment. (left) Dischmabach catchment, CNRM-CM5–CLM4-8-17; (center) Guerbe catchment, CNRM-CM5–RCA4; and (right) Breggia catchment, HadGEM2-ES–RCA4. All figures are for the period 1980–2009. Note the different axis scale for the three discharge plots in (g)–(i).

After QM, discharge ($Q_{qm}$) tends to resemble $Q_{ref}$ more than $Q_{obs}$ (Fig. 8). This pattern is due to the calibration of HBV that uses the MeteoSwiss gridded product as forcing data and the QM of GCM–RCM data that uses the same MeteoSwiss data ($P_{obs}$ and $T_{obs}$) as reference for the bias correction. Figure 8a shows discharge for Dischmabach catchment, which is a catchment where $P_{qm}$ and $T_{qm}$ fit the annual cycle generally well (Figs. 6a,d). Because of the improvements in the representation of precipitation throughout the annual cycle, discharge is greatly improved; note that $Q_{qm}$ resembles $Q_{ref}$ more so than $Q_{obs}$ (Fig. 8a; section 4c). Discharge for the Breggia catchment (Fig. 8c) has a precipitation cycle that peaks twice within the annual year (Fig. 6c). QM improves the GCM–RCM precipitation cycle ($P_{qm}$), although negative biases remain in the summer months. The improvements of discharge were substantial for the Guerbe catchment. The Guerbe catchment has small differences in the annual cycle of precipitation, which is relatively difficult for an annual QM method to improve. The improvements seen in discharge are a testament to the seasonal bias correction performed. The Breggia and Dischmabach catchments demonstrate the tendency for $Q_{qm}$ to resemble $Q_{ref}$ rather than $Q_{obs}$. For the Guerbe catchment, $Q_{obs}$ and $Q_{ref}$ are very similar, and thus $Q_{qm}$ resembles both.

### b. *Evaluating overall effect of quantile mapping*

After exploring the impacts of QM for individual basins, we analyzed the impacts of QM in all catchments. Figure 9 shows whether QM leads to an improvement

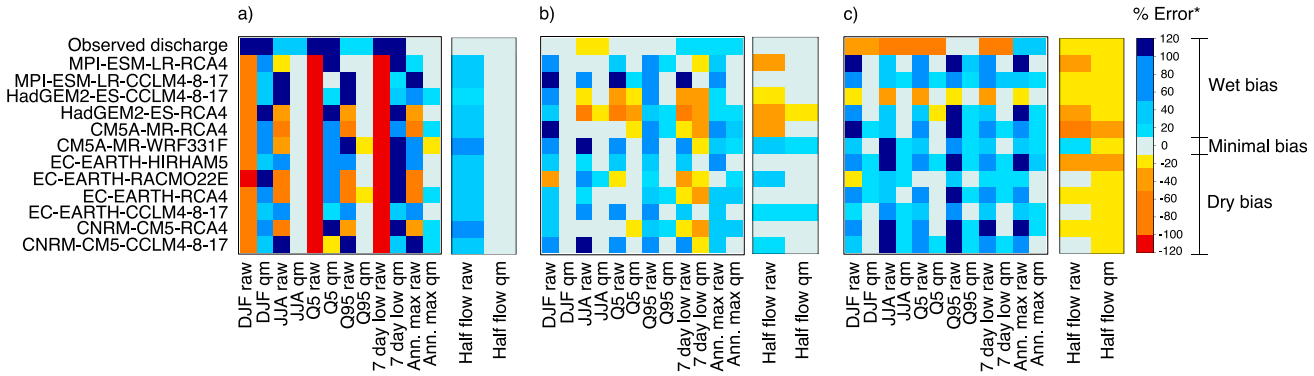



FIG. 7. Three example catchments are shown: (a) Dischmabach, (b) Guerbe, and (c) Breggia to demonstrate the overall impact of quantile mapping as well as the range in performance of the GCM–RCMs (*y* axis) according to hydrological metrics (*x* axis). The colors in the larger heat maps illustrate the values of $E_{obs}$, $E_{raw}$, or $E_{qm}$ (values in percent error). The colors in the smaller (half flow) heat maps are in units of days (see the * in the color bar legend). The top row within the heat plots shows observed discharge for comparative purposes. Observed discharge was not quantile mapped, thus the raw and quantile mapped columns are the same for this row.

(warm colors) or degradation (cool colors) of the hydrological simulations (i.e., it shows the difference between the absolute value of $E_{raw}$ and absolute value of $E_{qm}$ for each variable and climate model). The overall color pattern is predominantly warm tones, which implies that QM has a generally beneficial impact on discharge metrics. In 91% of all instances (all GCM–RCMs, catchments, and metrics considered), QM was found to improve discharge. By separating out high-flow metrics (JJA, Q95, annual maximum) from low-flow metrics (DJF, Q5, 7-day low flow), we found that high-flow metrics show generally greater improvement (96% improvement rate) from quantile mapping compared to low-flow metrics (87% improvement rate). Although low-flow metrics clearly did not improve as much as high-flow metrics, the initial calculation of $E_{raw}$ and $E_{qm}$ was very sensitive, especially when small values were compared to one another. Therefore, discharge can overall be greatly improved after QM, while low-flow metrics still show degradation, as, for instance, in the case in the Dischmabach catchment in Fig. 9.

### c. *Ranking climate models*

To synthesize our results, raw and quantile mapped GCM–RCMs were ranked according to the performance of runoff simulations, which are based on the precipitation and temperature extracted from the GCM–RCMs. To synthesize our results, we combined all of the hydrological variables into a single metric, referred to as "All metrics." The calculation of All metrics entails taking the median across all of the hydrological metrics (besides the half-flow date, which has a different unit) and all of the catchments for a particular GCM–RCM. The median was chosen in order to prevent the ranking from being overly affected by a particularly poor performing metric or catchment (e.g., low-flow metric).

Figure 10 shows the ranking based on $E_{raw}$ (Fig. 10a) and $E_{qm}$ (Fig. 10b), where $E_{raw}$ and $E_{qm}$ represent the median of all catchments. Observed discharge is also shown in the ranking for reference, based on $E_{obs}$. The order of the GCM–RCMs along the *y* axis was determined based on the ranking of the All metrics column. Within

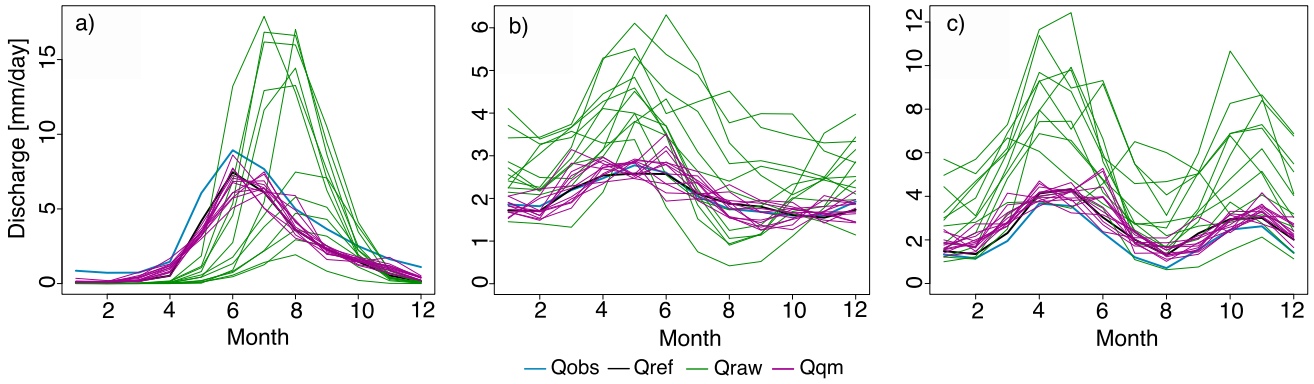



FIG. 8. Discharge for the (a) Dischmabach, (b) Guerbe, and (c) Breggia catchments with all GCM–RCMs shown for the period 1980–2009.

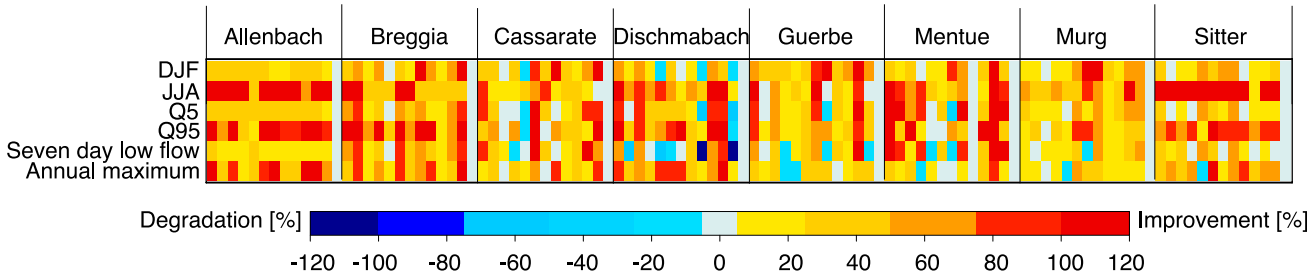



FIG. 9. Heat plot showing the difference (in absolute value) between the relative errors of discharge $|E_{raw} - E_{qm}|$ according to various metrics. The columns correspond to the GCM-RCM simulations in the order shown in Fig. 7 [e.g., the first column within the Allenbach section corresponds to the GCM-RCM (CNRM-CM5-CCLM4-8-17), and the last column within the Allenbach section corresponds to "Observed discharge"].

Fig. 10a, observed discharge ranks high in comparison to the GCM–RCMs, which is in strong contrast to Fig. 10b, where observed discharge ranks last. The switch in placement of observed discharge is due to the general improvement of the GCM–RCM performance after quantile mapping. It is especially noteworthy that within Fig. 10b, the ranking of observed discharge is worse than any GCM–RCM forcing. The result in Fig. 10b shows that after QM, the percent error between $Q_{qm}$ and $Q_{ref}$ is smaller than the percent error between $Q_{obs}$ and $Q_{ref}$. This pattern is caused by both the bias correction of the GCM–RCM temperature and precipitation as well as the calibration of HBV since the calibration of HBV was done so that $Q_{ref}$ should resemble $Q_{obs}$ (see section 4c for more discussion). The rank of $Q_{obs}$ as last compared to bias-corrected GCM–RCMs is a confirmation of the ability of QM to improve discharge metrics. In addition, GCM–RCMs also change their rank between Figs. 10a and 10b, despite the uniform application of QM. This is in part because the differences between the bias-corrected GCM–RCMs are reduced and a single percent error can change the order of ranking. Results show there is no general pattern pointing to a decidedly single-best GCM or RCM.

Besides noting the performance order of GCM–RCMs as seen in Fig. 10, it is also important to show the amount of improvement (in percent error) one would achieve if choosing between the top and the lowest-ranked GCM–RCM or between QM and raw GCM–RCM data. Figure 11 shows a bar graph comparing quantile mapped discharge data ($E_{qm}$) to raw ($E_{raw}$) discharge data for All metrics. Raw GCM–RCMs show more variability with percent errors ranging from 26% to 88%. Quantile mapped GCM–RCMs range from 4% to 11%. QM clearly reduces differences between $Q_{ref}$ and $Q_{qm}$. Note that the reduction in overall bias also causes discharge stemming from different GCM–RCM forcings to resemble one another (see section 4c for more discussion). Figure 11 demonstrates this result, where the "All metrics quantile mapped" color bars show similar levels of percent error.

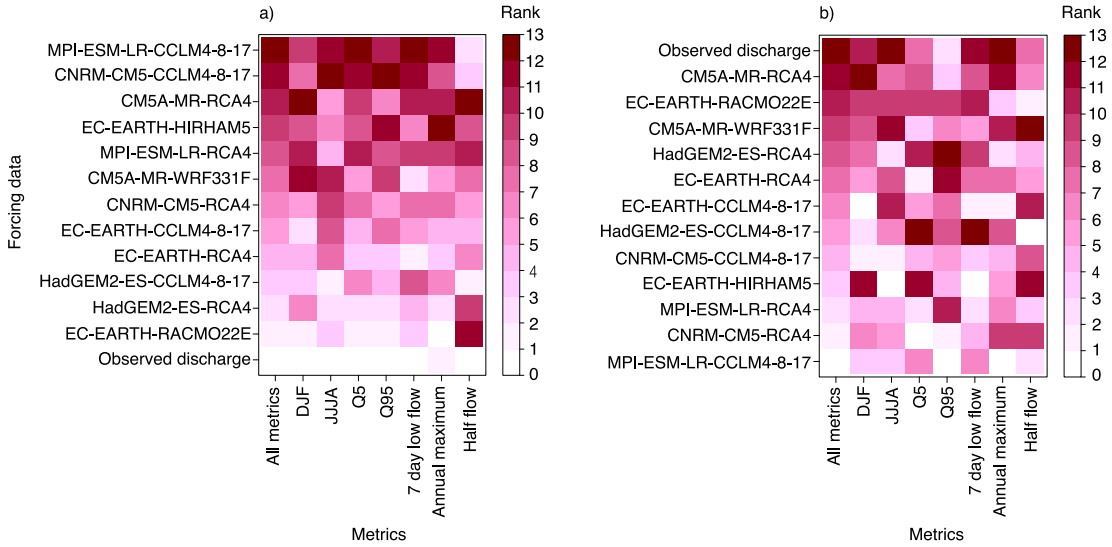



FIG. 10. (a) Raw and (b) QM GCM–RCMs ($y$ axis) ranked according to their performance for various hydrological metrics ($x$ axis) across all catchments. The placement of the GCM–RCMs along the $y$ axis is determined by their rank within the All metrics column. Note that observed discharge ranks high in (a) and ranks low in (b).

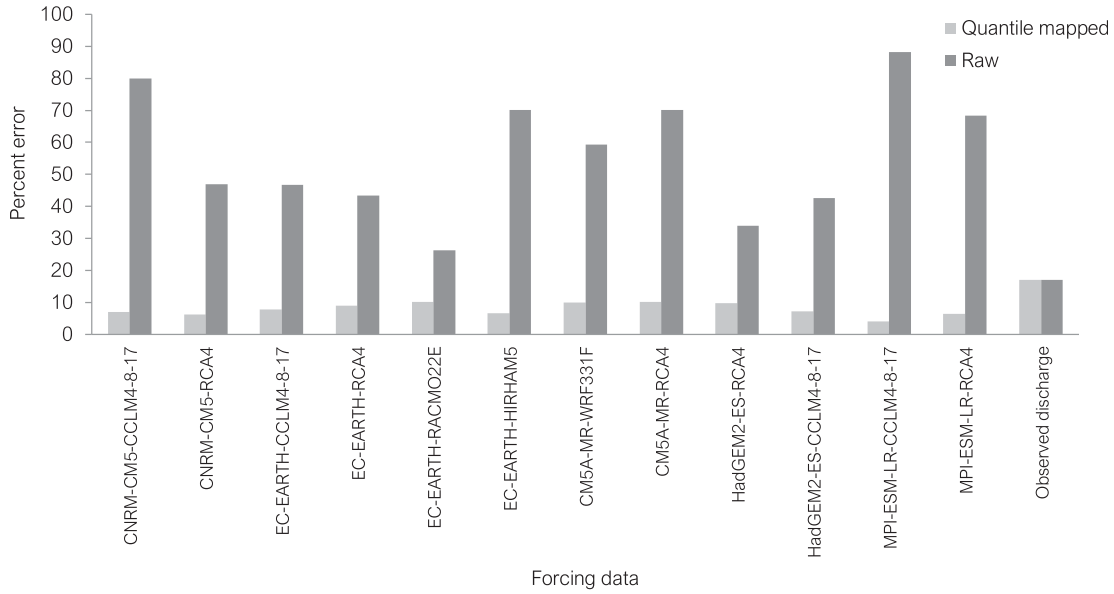



FIG. 11. Raw (dark gray) and QM (light gray) GCM–RCMs (*x* axis) and their median performance across all hydrological metrics (i.e., All metrics) and all catchments (*y* axis).

## 4. Discussion

### a. *How do RCM biases impact the representation of hydrological processes?*

Biases in EURO-CORDEX data can be significant at the catchment scale and can have substantial effects on the simulated discharge. Our study identified wet precipitation biases, occurring in the winter/spring months with occasional dry biases in the summer months, as well as generally cold temperature biases, especially at high elevations. Previous studies such as Frei et al. (2018), who examined precipitation from EURO-CORDEX RCMs over the Alpine region, found that prior to bias correction, snowfall amounts at high elevations can be considerably overestimated. Wet precipitation biases over Switzerland have also been found in earlier GCM–RCM model generations (ENSEMBLES; van der Linden et al. 2009) as shown by Fischer et al. (2012) and Addor et al. (2016). Our study identified that, within high-elevation catchments, the combination of excessive precipitation with a cold bias translates into greater discharge values and delays in spring melt. Temperature biases are not as strong in low- to mid-elevation catchments and thus, in these catchments, the timing of discharge (using raw GCM–RCM data to force HBV) resembles that of $Q_{obs}$ and $Q_{ref}$ relatively closely.

Our results confirm the general beneficial use of quantile mapping, which has been reported in previous studies (Themeßl et al. 2011a; Teutschbein and Seibert 2012; Räisänen and Räty 2013). However, after the application of a seasonal QM, biases are often still present, although reduced. This is not surprising, as previous literature has pointed out. Addor and Seibert (2014) show that after performing a bias correction of precipitation over a daily time step, for instance, discrepancies between the observations and the GCM–RCM simulations can remain for other time scales. Our work shows a similar manifestation of this concept in that daily bias-corrected precipitation and temperature data contain biases on the monthly time scale (Figs. 6a–f). In addition, the discharge metrics used herein (Figs. 7a–c) are sensitive to various time scales. Discharge itself is the end result of processes covering a wide range of time scales. Therefore, it can be expected that discharge biases can remain even after a seasonal bias correction has been applied to GCM–RCM temperature and precipitation. Other instances where QM did not lead to an improvement in discharge simulations (see Fig. 7a, DJF qm or Q5 qm columns) occur when biases were not significant to start with. In particular, biases in the low-flow period in high-elevation catchments can increase after QM, but the associated volume of water is typically low. In other words, although they can be large relative biases, they are not necessarily significant in absolute terms. Overall, both the magnitude and timing of discharge are improved (i.e., QM causes $Q_{qm}$ to more closely resemble $Q_{obs}$ and $Q_{ref}$).

There are of course limitations related to any bias correction method. Within this study, a univariate bias correction was used, which means that temperature and precipitation were corrected independently of each other. This method is limited in that it does not specifically consider the intervariable dependence structure between temperature and precipitation. In addition, we

applied quantile mapping at the daily time scale, although biases also exist over other time scales (e.g., decadal, subdaily). More advanced methods exist to accommodate for these factors, such as the multivariate recursive quantile nesting bias correction (MRQNBC; Mehrotra and Sharma 2016), which corrects individual variable attributes that lead to correction of dependence biases between multiple variables. This method also corrects for lag-1 dependence and cross-dependence attributes over multiple time scales. Another promising method to correct for biases on multiple time scales is the frequency bias correction method (FBC; Nguyen et al. 2016), which corrects for biases in the frequency domain. While other bias correction methods exist, the goal of this study was not to compare or advocate for a particular bias correction method. Rather, we demonstrate that the utilization of hydrological modeling can be used to evaluate climate simulations and assess whether a bias correction technique was successful at reducing the biases relevant for hydrological impact studies.

### b. *How does quantile mapping deal with the snow towers built by some climate models?*

Some GCM–RCM biases clearly indicate that the simulations are physically unrealistic (e.g., snow depth of over 200 m at the end of the historical simulation ending in 2005). The question then arises whether it is meaningful to perform bias correction or whether the model should be excluded from the ensemble. For instance, the excessive buildup of GCM–RCM snow depth within some simulations provides a testing ground to investigate the sensitivity of discharge to input derived from snow tower affected simulations. Terzago et al. (2017) explored SWE in the Alps and used various observational SWE datasets to evaluate CMIP5 GCM and EURO-CORDEX RCM simulations. Their study identified some extremely high values of SWE, originating from excessive accumulation of snow. They chose to eliminate these climate models from the rest of their analysis. Besides the particular models that build snow towers, Terzago et al. (2017) report that all RCMs simulate more SWE, along mountain ridges, than any of the reference datasets they considered. They partly attribute this bias to the higher resolution of the RCMs compared to the resolution of the reference datasets. Higher resolution allows for better representation of heterogeneous mountain topography and therefore for colder temperatures at high elevation. In addition, the large SWE values from the RCMs can also be explained by cold (e.g., RACMO22E) and wet biases (HIRHAM5) in relation to observations. Figure 10 shows that the ranking for all catchments considered, a GCM–RCM with a snow tower (EC-EARTH–RACMO22E) ranks high in comparison to other GCM–RCMs. However, most catchments do not have snow towers associated with them. When considering a single catchment that has a snow tower (e.g., Allenbach catchment), the EC-EARTH–RACMO22E also ranks high. This simple analysis shows that the presence of a snow tower within a GCM–RCM does not necessarily affect temperature and precipitation (over the historical period) to the point that resulting streamflow simulations can detect the presence of a snow tower. However, the feedback between a snow tower and temperature over the future period has been shown to further reduce the climate change signal (Frei et al. 2018). Therefore, in the case of snow tower–affected GCM–RCMs, the performance of a GCM–RCM over both the historical period and the future should be considered when deciding whether a model is viable for use or not. The brief analysis of snow tower–affected GCM–RCMs herein points to the need for a greater dialogue regarding which types of biases should warrant inclusion/exclusion within a climate model ensemble.

### c. *The value of GCM–RCM ranking using hydrological modeling*

Hydrological modeling allows for a combined assessment of the hydrologically important aspects of precipitation and temperature time series. This study provides a ranking that simultaneously considers a multitude of factors relevant for hydrological modeling (Fig. 10). The value of such a ranked set of GCM–RCMs strongly depends on the intended use of the ranking.

For the evaluation of climate model simulations, ranking raw GCM–RCMs according to hydrological performance provides a new perspective on climate model realism. By considering discharge metrics as a standard from which to rank and assess the performance of climate models, one can automatically account for interactions between atmospheric variables ($P$ and $T$) at the catchment scale, including both linear and nonlinear hydrological processes.

For the selection of climate models for impact modeling, it is important to underline that QM largely improves the simulated discharge from a hydrological model. We found that across all GCM–RCMs, catchments, and discharge metrics, QM led to improvement of the simulations in 91% of the cases. QM simulations should therefore be considered as more reliable by end users (e.g., water managers) than hydrological projections driven by raw climate simulations. Another key result of this study is that QM causes the streamflow simulations to converge (see, e.g., Fig. 8). As Fig. 10b shows, after QM, all GCM–RCMs perform better than $Q_{obs}$ (relative to $Q_{ref}$). In other words, the

difference between any QM GCM–RCM ($Q_{qm}$) and $Q_{ref}$ is smaller than the difference between $Q_{obs}$ and $Q_{ref}$. This is because QM corrects GCM–RCM output so that it matches observational data. When quantile mapped GCM–RCM simulations ($P_{qm}$ and $T_{qm}$) were used to force HBV, it results in improved representation of hydrological processes. Also, $Q_{ref}$ and $Q_{obs}$ may not be similar enough, implying that the structure and/or the parameter sets of HBV need further tuning. The level of calibration and the climate model culling standards should be determined by the needs of the user. It is important to note that we assess climate models under current climate and do not consider their skill under future climate. Under future conditions, the spread among the bias-corrected simulations will be greater than under current conditions and should be accounted for when selecting GCM–RCMs. Finally, climate model selection should also be informed by the errors in hydrological metrics most relevant to the end users.

It is important to recognize that although the observed precipitation dataset (RHiresD) has a nominal resolution of about 2 km, its effective resolution is significantly coarser at 15–20 km. RHiresD is the finest precipitation dataset currently available over Switzerland, yet the catchment sizes are subgrid to its effective grid, meaning that RHiresD does not fully resolve the precipitation events in our study catchments. An implication is that catchment-scale precipitation estimates are more uncertain than what might be inferred based on the 2-km grid RHiresD data. These uncertainties can influence the ranking of the GCM–RCMs, since in some cases, GCM–RCMs may capture precipitation more realistically than RHiresD and with our setup, these GCM–RCMs would be penalized (Gómez-Navarro et al. 2012; Addor and Fischer 2015; Prein and Gobiet 2017). Another issue arises given that the study catchment sizes are all smaller than one entire RCM grid cell. This makes the evaluation and ranking of raw GCM–RCM output challenging since these models were not designed to represent features at the spatial scale at which they are being evaluated. To help alleviate this issue, the effects of elevation on temperature and precipitation were explicitly accounted for prior to the ranking of the raw GCM–RCM data. For future work, a way to overcome these issues would be to work with larger catchments, but then the risk of perturbation of the hydrological time series because of human interventions would be higher than in the research catchments considered here.

## 5. Conclusions

This study investigated how biases in EURO-CORDEX GCM–RCM simulations impact the representation of hydrological processes. Quantile mapping (QM) was shown to be highly effective in improving discharge metrics. When all catchments, streamflow metrics and GCM–RCMs are considered, QM leads to an improvement in the vast majority (91%) of cases. When inspecting the annual discharge cycle, it is clear that QM overall improves the simulated discharge, often because of the more realistic simulation of snow-related processes. Most of the occasional degradations are observed in low-flow metrics. These degradations may be large in relative terms, but they are typically small when compared to the improvements over the rest of the discharge cycle.

Our study demonstrates that hydrological modeling can be used to evaluate and rank climate model simulations in an integrated way at the catchment scale. For climate modelers, it is a way to gain novel insights into climate model realism. For impact modelers, who have to select climate models for hydrologic modeling, this evaluation approach is a way to assess the sensitivity of hydrological simulations to known biases, such as the existence of snow towers in some EURO-CORDEX simulations.

Another key finding of this study is that applying QM causes the convergence of hydrological simulations driven by GCM–RCMs under current climate. This stems from the use of a common reference observational dataset for the bias correction of the simulated atmospheric forcing. Since the cumulative distribution of GCM–RCM temperature and precipitation is forced to mimic that of the observations, it causes the resulting hydrological simulations to resemble each other under current climate. This implies that the ranking of GCM–RCM simulations after QM provides limited insights. Rather, ranking GCM–RCMs prior to bias correction is recommended, especially when performed over catchments that are large in comparison to the resolution of the RCM grid. At this stage, it is unclear whether QM will cause a convergence of the future climate change impacts on discharge. Next steps include the application of this analysis to operational decision-making, which will include the consideration of future climate change impacts on hydrology.

The use of hydrological modeling to assess the performance of climate models has received little attention so far. Our combination of the newest generation of GCM–RCM simulations within a hydrological framework allows for the simultaneous consideration of a wide range of climate models, hydrologic regimes, and streamflow variables. Hydrological modeling provides new insights to climate modelers and end users and represents a novel way to assess the realism and support the selection of climate models.

*Acknowledgments.* This study was funded by the Swiss National Science Foundation via the SNF Grant 200020_156606: Hydrological climate change impact

assessment—Addressing the uncertainties (HIMAUI). We thank Marc Vis for support related to HBV, Sven Kotlarski for providing useful insights into CORDEX data, and the two anonymous reviewers for their comments. We acknowledge the Swiss Federal Offices for the Environment (FOEN), the Meteorology and Climatology Offices of Switzerland (MeteoSwiss), and the European Coordinated Regional Downscaling Experiment (EURO-CORDEX) for the hydrological, atmospheric, and climate modeled data, respectively.

## REFERENCES

Addor, N., and J. Seibert, 2014: Bias correction for hydrological impact studies - beyond the daily perspective. *Hydrol. Processes*, **28**, 4823–4828, https://doi.org/10.1002/hyp.10238.

——, and E. M. Fischer, 2015: The influence of natural variability and interpolation errors on bias characterization in RCM simulations. *J. Geophys. Res. Atmos.*, **120**, 10 180–10 195, https://doi.org/10.1002/2014JD022824.

——, M. Rohrer, R. Furrer, and J. Seibert, 2016: Propagation of biases in climate models from the synoptic to the regional scale: Implications for bias adjustment. *J. Geophys. Res. Atmos.*, **121**, 2075–2089, https://doi.org/10.1002/2015JD024040.

Bergström, S., 1976: *Development and Application of a Conceptual Runoff Model for Scandinavian Catchments.* Norrköping, 134 pp.

Criss, R. E., and W. E. Winston, 2008: Do Nash values have value? Discussion and alternate proposals. *Hydrol. Processes*, **22**, 2723–2725, https://doi.org/10.1002/hyp.7072.

Fischer, A. M., A. P. Weigel, C. M. Buser, R. Knutti, H. R. Künsch, M. A. Liniger, C. Schär, and C. Appenzeller, 2012: Climate change projections for Switzerland based on a Bayesian multi-model approach. *Int. J. Climatol.*, **32**, 2348–2371, https://doi.org/10.1002/joc.3396.

Fowler, H. J., S. Blenkinsop, and C. Tebaldi, 2007: Linking climate change modelling to impacts studies: Recent advances in downscaling techniques for hydrological modelling. *Int. J. Climatol.*, **27**, 1547–1578, https://doi.org/10.1002/joc.1556.

Frei, C., 2014: Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances. *Int. J. Climatol.*, **34**, 1585–1605, https://doi.org/10.1002/joc.3786.

——, and C. Schär, 1998: A precipitation climatology of the Alps from high-resolution rain-gauge observations. *Int. J. Climatol.*, **18**, 873–900, https://doi.org/10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9.

Frei, P., S. Kotlarski, M. A. Liniger, and C. Schär, 2018: Future snowfall in the Alps: Projections based on the EURO-CORDEX regional climate models. *Cryosphere*, **12**, 1–24, https://doi.org/10.5194/tc-12-1-2018.

Gleckler, P. J., K. E. Taylor, and C. Doutriaux, 2008: Performance metrics for climate models. *J. Geophys. Res.*, **113**, D06104, https://doi.org/10.1029/2007JD008972.

Gómez-Navarro, J. J., J. P. Montvez, S. Jerez, P. Jiménez-Guerrero, and E. Zorita, 2012: What is the role of the observational dataset in the evaluation and scoring of climate models? *Geophys. Res. Lett.*, **39**, L24701, https://doi.org/10.1029/2012GL054206.

Gudmundsson, L., 2016: Qmap: Statistical transformations for post-processing climate model output. R package, version 1.0-4, https://cran.r-project.org/web/packages/qmap/index.html.

——, J. B. Bremnes, J. E. Haugen, and T. Engen-Skaugen, 2012: Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations—A comparison of methods. *Hydrol. Earth Syst. Sci.*, **16**, 3383–3390, https://doi.org/10.5194/hess-16-3383-2012.

Johnson, F., and A. Sharma, 2015: What are the impacts of bias correction on future drought projections? *J. Hydrol.*, **525**, 472–485, https://doi.org/10.1016/j.jhydrol.2015.04.002.

Jury, M. W., A. F. Prein, H. Truhetz, and A. Gobiet, 2015: Evaluation of CMIP5 models in the context of dynamical downscaling over Europe. *J. Climate*, **28**, 5575–5582, https://doi.org/10.1175/JCLI-D-14-00430.1.

Kotlarski, S., and Coauthors, 2014: Regional climate modeling on European scales: A joint standard evaluation of the EURO-CORDEX RCM ensemble. *Geosci. Model Dev.*, **7**, 1297–1333, https://doi.org/10.5194/gmd-7-1297-2014.

Li, C., E. Sinha, D. E. Horton, N. S. Diffenbaugh, and A. M. Michalak, 2014: Joint bias correction of temperature and precipitation in climate model simulations. *J. Geophys. Res. Atmos.*, **119**, 13 153–13 162, https://doi.org/10.1002/2014JD022514.

Lindström, G., B. Johansson, M. Persson, M. Gardelin, and S. Bergström, 1997: Development and test of the distributed HBV-96 hydrological model. *J. Hydrol.*, **201**, 272–288, https://doi.org/10.1016/S0022-1694(97)00041-3.

Maraun, D., 2013: Bias correction, quantile mapping, and downscaling: Revisiting the inflation issue. *J. Climate*, **26**, 2137–2143, https://doi.org/10.1175/JCLI-D-12-00821.1.

Mehrotra, R., and A. Sharma, 2016: A multivariate quantile-matching bias correction approach with auto- and cross-dependence across multiple time scales: Implications for downscaling. *J. Climate*, **29**, 3519–3539, https://doi.org/10.1175/JCLI-D-15-0356.1.

Mendlik, T., and A. Gobiet, 2016: Selecting climate simulations for impact studies based on multivariate patterns of climate change. *Climatic Change*, **135**, 381–393, https://doi.org/10.1007/s10584-015-1582-0.

MeteoSwiss, 2010: SwissMetNet: Ein Messnetz für die Zukunft. MeteoSwiss, https://naturwissenschaften.ch/service/news/77197-swissmetnet—ein-messnetz-fuer-die-zukunft.

——, 2013a: Documentation of MeteoSwiss grid-data products: Daily precipitation (final analysis): RhiresD. MeteoSwiss, 4 pp., https://www.meteoswiss.admin.ch/content/dam/meteoswiss/de/service-und-publikationen/produkt/raeumliche-daten-niederschlag/doc/ProdDoc_RhiresD.pdf.

——, 2013b: Daily mean, minimum and maximum temperature: TabsD, TminD, TmaxD. MeteoSwiss, 4 pp., https://www.meteoswiss.admin.ch/content/dam/meteoswiss/de/service-und-publikationen/produkt/raeumliche-daten-temperatur/doc/ProdDoc_TabsD.pdf.

Nash, J. E., and J. V. Sutcliffe, 1970: River flow forecasting through conceptual models part I — A discussion of principles. *J. Hydrol.*, **10**, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6.

Nguyen, H., R. Mehrotra, and A. Sharma, 2016: Correcting for systematic biases in GCM simulations in the frequency domain. *J. Hydrol.*, **538**, 117–126, https://doi.org/10.1016/j.jhydrol.2016.04.018.

Pechlivanidis, I. G., J. Olsson, D. Sharma, T. Bosshard, and K. C. Sharma, 2015: Assessment of the climate change impacts on the water resources of the Luni Region, India. *Global NEST J.*, **17**, 29–40, https://doi.org/10.30955/gnj.001370.

Prein, A. F., and A. Gobiet, 2017: Impacts of uncertainties in European gridded precipitation observations on regional

climate analysis. *Int. J. Climatol.*, **37**, 305–327, https://doi.org/10.1002/joc.4706.

Räisänen, J., and O. Räty, 2013: Projections of daily mean temperature variability in the future: Cross-validation tests with ENSEMBLES regional climate simulations. *Climate Dyn.*, **41**, 1553–1568, https://doi.org/10.1007/s00382-012-1515-9.

Reichler, T., and J. Kim, 2008: How well do coupled models simulate today's climate? *Bull. Amer. Meteor. Soc.*, **89**, 303–311, https://doi.org/10.1175/BAMS-89-3-303.

Salathé, E. P., P. W. Mote, and M. W. Wiley, 2007: Review of scenario selection and downscaling methods for the assessment of climate change impacts on hydrology in the United States Pacific Northwest. *Int. J. Climatol.*, **27**, 1611–1621, https://doi.org/10.1002/joc.1540.

Schaefli, B., B. Hingray, and A. Musy, 2007: Climate change and hydropower production in the Swiss Alps: Quantification of potential impacts and related modelling uncertainties. *Hydrol. Earth Syst. Sci.*, **11**, 1191–1205, https://doi.org/10.5194/hess-11-1191-2007.

Schwarb, M., 2000: The alpine precipitation climate: Evaluation of a high-resolution analysis scheme using comprehensive rain-gauge data. Ph.D. dissertation, Swiss Federal Institute of Technology, 131 pp.

Seibert, J., 2000: Multi-criteria calibration of a conceptual runoff model using a genetic algorithm. *Hydrol. Earth Syst. Sci.*, **4**, 215–224, https://doi.org/10.5194/hess-4-215-2000.

——, and M. J. P. Vis, 2012: Teaching hydrological modeling with a user-friendly catchment-runoff-model software package. *Hydrol. Earth Syst. Sci.*, **16**, 3315–3325, https://doi.org/10.5194/hess-16-3315-2012.

Terzago, S., J. Von Hardenberg, E. Palazzi, and A. Provenzale, 2017: Snow water equivalent in the Alps as seen by gridded data sets, CMIP5 and CORDEX climate models. *Cryosphere*, **11**, 1625–1645, https://doi.org/10.5194/tc-11-1625-2017.

Teutschbein, C., and J. Seibert, 2012: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods. *J. Hydrol.*, **456–457**, 12–29, https://doi.org/10.1016/j.jhydrol.2012.05.052.

Themeßl, M. J., G. Andreas, and A. Leuprecht, 2011a: Empirical-statistical downscaling and error correction of daily precipitation from regional climate models. *Int. J. Climatol.*, **31**, 1530–1544, https://doi.org/10.1002/joc.2168.

——, A. Gobiet, and G. Heinrich, 2011b: Empirical-statistical downscaling and error correction of regional climate models and its impact on the climate change signal. *Climatic Change*, **112**, 449–468, https://doi.org/10.1007/s10584-011-0224-4.

van der Linden, P., J. F. B. Mitchell, and P. Gilbert, 2009: ENSEMBLES: Climate change and its impacts: Summary of research and results from the ENSEMBLES project. P. Van Der Linden and J. F. B. Mitchell, Eds., Met Office Hadley Centre, 160 pp.

Wilcke, R. A. I., and L. Bärring, 2016: Selecting regional climate scenarios for impact modelling studies. *Environ. Modell. Software*, **78**, 191–201, https://doi.org/10.1016/j.envsoft.2016.01.002.

——, T. Mendlik, and A. Gobiet, 2013: Multi-variable error correction of regional climate models. *Climatic Change*, **120**, 871–887, https://doi.org/10.1007/s10584-013-0845-x.

Zehe, E., and M. Sivapalan, 2009: Threshold behaviour in hydrological systems as (human) geo-ecosystems: Manifestations, controls, implications. *Hydrol. Earth Syst. Sci.*, **13**, 1273–1297, https://doi.org/10.5194/hess-13-1273-2009.

PAPER III

Hydrol. Earth Syst. Sci., 23, 1339–1354, 2019
https://doi.org/10.5194/hess-23-1339-2019


Hydrology and Earth System Sciences
Open Access
EGU

# Effects of univariate and multivariate bias correction on hydrological impact projections in alpine catchments

**Judith Meyer**[1,a], **Irene Kohn**[1], **Kerstin Stahl**[1], **Kirsti Hakala**[2], **Jan Seibert**[2,3], **and Alex J. Cannon**[4]

[1]Faculty of Environment and Natural Resources, University of Freiburg, 79098 Freiburg, Germany
[2]Department of Geography, University of Zurich, 8057 Zurich, Switzerland
[3]Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, Uppsala, Sweden
[4]Climate Research Division, Environment and Climate Change Canada, BC V8W 2Y2, Victoria, Canada
[a]now at: Catchment and Eco-Hydrology Research Group, Luxembourg Institute of Science and Technology, 4362 Esch-sur-Alzette, Luxembourg

**Correspondence:** Irene Kohn (irene.kohn@hydrology.uni-freiburg.de)

Received: 6 June 2018 – Discussion started: 28 June 2018
Revised: 22 February 2019 – Accepted: 25 February 2019 – Published: 11 March 2019

**Abstract.** Alpine catchments show a high sensitivity to climate variation as they include the elevation range of the snow line. Therefore, the correct representation of climate variables and their interdependence is crucial when describing or predicting hydrological processes. When using climate model simulations in hydrological impact studies, forcing meteorological data are usually downscaled and bias corrected, most often by univariate approaches such as quantile mapping of individual variables, neglecting the relationships that exist between climate variables. In this study we test the hypothesis that the explicit consideration of the relation between air temperature and precipitation will affect hydrological impact modelling in a snow-dominated mountain environment. Glacio-hydrological simulations were performed for two partly glacierized alpine catchments using a recently developed multivariate bias correction method to post-process EURO-CORDEX regional climate model outputs between 1976 and 2099. These simulations were compared to those obtained by using the common univariate quantile mapping for bias correction. As both methods correct each climate variable's distribution in the same way, the marginal distributions of the individual variables show no differences. Yet, regarding the interdependence of precipitation and air temperature, clear differences are notable in the studied catchments. Simultaneous correction based on the multivariate approach led to more precipitation below air temperatures of 0 °C and therefore more simulated snowfall than with the data of the univariate approach. This difference translated to considerable consequences for the hydrological responses of the catchments. The multivariate bias-correction-forced simulations showed distinctly different results for projected snow cover characteristics, snowmelt-driven streamflow components, and expected glacier disappearance dates. In all aspects – the fraction of precipitation above and below 0 °C, the simulated snow water equivalents, glacier volumes, and the streamflow regime – simulations resulting from the multivariate-corrected data corresponded better with reference data than the results of univariate bias correction. Differences in simulated total streamflow due to the different bias correction approaches may be considered negligible given the generally large spread of the projections, but systematic differences in the seasonally delayed streamflow components from snowmelt in particular will matter from a planning perspective. While this study does not allow conclusive evidence that multivariate bias correction approaches are generally preferable, it clearly demonstrates that incorporating or ignoring inter-variable relationships between air temperature and precipitation data can impact the conclusions drawn in hydrological climate change impact studies in snow-dominated environments.

## 1 Introduction

With global change, hydrological processes in high elevation regions have been significantly impacted (Messerli et

al., 2004). In the European Alps, the observed increase in air temperature is a trend that is expected to continue in the future. Future precipitation changes are less clear, with an expected slight increase in winter precipitation (Gobiet et al., 2014; Kotlarski et al., 2016). The hydrology of alpine catchments is especially sensitive to these changing climate variables (Köplin et al., 2010). High elevations in the Alps are still characterized by snow cover and the existence of glaciers. However, rising air temperatures and a consequent upward shift of the zero-degree isotherm has led to a decrease in snow accumulation and an increase in glacier melt (Pellicciotti et al., 2010). Due to shrinking glacier areas, the glacial influence in these streamflow regimes has decreased. This is especially notable during late summer when water from ice melt can constitute a notable percentage of total streamflow. With progressive glacier retreat, the ice melt contribution to streamflow is expected to decrease (Jansson et al., 2003; Hock, 2005; Moore et al., 2009; Huss and Hock, 2018). The interdependence of air temperature and precipitation is particularly important for hydrological systems as it determines the physical state of precipitation. Bosshard et al. (2014) showed that an air-temperature-dependent shift from snowfall to rain has notable effects on catchment water storage and seasonal water availability in such an environment. A correct representation of climate variables and their interdependence is therefore essential in hydrological simulations of glacierized catchments.

In hydrological climate change impact studies, post-processing of climate model data has become a standard procedure. Despite continuous progress, raw outputs from regional climate models differ largely from observational reference data due to both spatial mismatches and systematic biases. Therefore, climate model outputs are downscaled and biases are adjusted statistically before being used in hydrological simulations (Ehret et al., 2012; Maraun, 2016; Teutschbein and Seibert, 2012). Many empirical statistical techniques have been developed to post-process climate model outputs for these purposes. For hydrological impact studies quantile mapping approaches, which correct for biases in the data's entire distribution, have often been recommended (Teutschbein and Seibert, 2012; Gudmundsson et al., 2012; Chen et al., 2013). However, these approaches correct the climate variables independently from one another. The interdependence of key climate variables, such as air temperature and precipitation, can be especially important when modelling snow-dominated catchments due to the aforementioned threshold effects of the transition of rain to snowfall or the conditions required for snowmelt and ice melt.

Studies that analysed inter-variable aspects of bias correction showed that univariate quantile mapping retains the inter-variable dependencies as represented by the raw climate model output data (Wilcke et al., 2013; Ivanov and Kotlarski, 2017). But these may not correspond to the local interdependencies in observations. To account for interdependencies, multivariate bias correction approaches have been developed that allow for the preservation of the interdependence of climate variables as represented by the target observation data throughout the bias correction process (Li et al., 2014; Cannon, 2016, 2018a; Mehrotra and Sharma, 2015, 2016). A correction procedure that preserves the climate variables' interdependence may be considered more appropriate for subsequent impact analyses, such as the application of a calibrated hydrological model using multiple variables, than univariate techniques that ignore biases in inter-variable relationships (Cannon, 2018a).

While many studies have evaluated bias correction methods in terms of their effects on the actual variables of precipitation and air temperature themselves, studies that use impact models to investigate the consequence of bias correction in the modelled impacts are still rare. So far, there have been only a few studies (Räty et al., 2018; Chen et al., 2018) that investigated the effect of using a multivariate bias correction technique on hydrological projections. Chen et al. (2018) found that jointly corrected precipitation and air temperature data better modelled eleven out of twelve catchments in the calibration period than the meteorological data that was corrected with a univariate method. An advantage of using a bivariate bias correction approach was not evident for the coldest snow-dominated catchment of the sample though. Hydrological simulations by Räty et al. (2018) generally did not substantially benefit from bivariate bias correction approaches, but when looking more specifically, simulations of high flows and snow water equivalents in snow-influenced catchments improved slightly.

In this study we investigate the hypothesis that the explicit consideration of the relation between air temperature and precipitation in bias correction will affect hydrological impact modelling in environments dominated by snowmelt and glacier melt. Here, dependencies are known to matter most as they have cumulative effects over a season through snow storage and at multi-year timescales through the glacier mass balance. The approach of this study was therefore to conduct climate impact modelling experiments that allow comparison of the effects of univariate and multivariate bias correction of precipitation and air temperature input on the hydrological change in alpine catchments. The model experiments were conducted for two meso-scale partly glacierized catchments in the Swiss Alps, for which snow accumulation, glacier mass balance, and streamflow were simulated from 1976 to 2099.

## 2 Study catchments and data

### 2.1 Study area

Two partly glacierized meso-scale catchments in the Swiss Alps, in the headwater of the Rhine River, were examined in this study: the Hinterrhein catchment and the larger Schwarze Lütschine catchment (Fig. 1, Table 1). Based on

**Table 1.** Catchment characteristics including glacier cover information.

| | Area | Elevation | | | Glacier cover* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | | | | 1973 | | 2003 | | 2010 | |
| | ($km^2$) | Mean | Min | Max | ($km^2$) | (%) | ($km^2$) | (%) | ($km^2$) | (%) |
| | | | (m a.s.l.) | | | | | | | |
| Hinterrhein | 53.9 | 2357 | 1587 | 3387 | 9.1 | 17.8 | 4.7 | 8.7 | 3.8 | 7.1 |
| Schwarze Lütschine | 179.9 | 2059 | 648 | 4086 | 37.0 | 23.5 | 34.4 | 19.1 | 29.7 | 16.5 |

* Based on glacier inventories by Müller et al. (1976) and Maisch et al. (2000) for 1973, Paul et al. (2011) for 2003, and Fischer et al. (2014) for 2010.

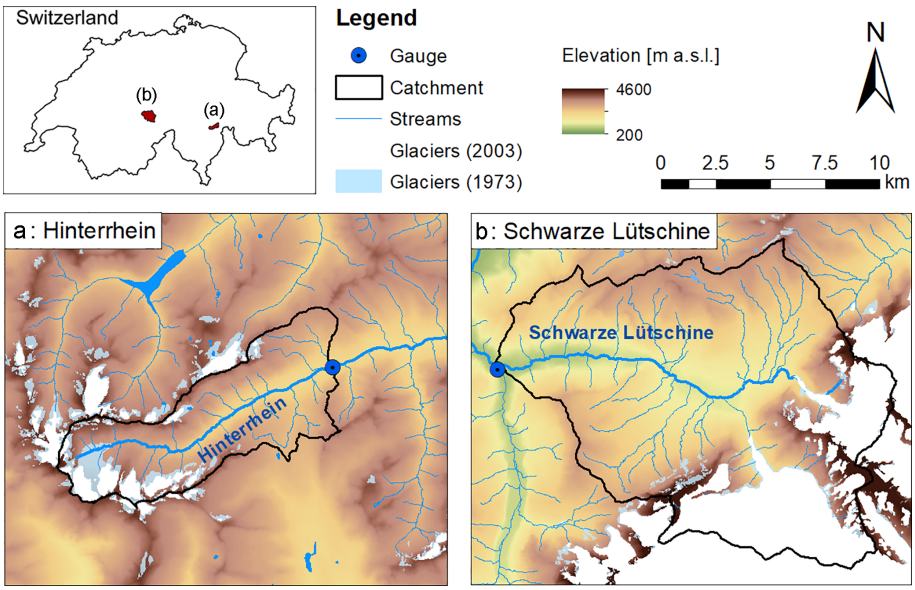



**Figure 1.** Map of the two study catchments and their location in Switzerland: Hinterrhein **(a)** and Schwarze Lütschine **(b)**.

the dataset by Freudiger et al. (2018), used in this study, around the year 1900 glacier coverage was approximately 32 % of the Hinterrhein catchment area and around 25 % of the Schwarze Lütschine catchment area. Glaciers in both catchments retreated considerably during the 20th century. The Hinterrhein catchment is characterized by small, scattered glaciers, which by 1973 lost around half their area, leading to a glacier coverage of only 7 % in 2010 (Table 1). In the Schwarze Lütschine catchment losses in relative glacier area have been smaller. This difference in glacier coverage is related to elevation with considerably higher maximum elevations in the Schwarze Lütschine catchment compared to the Hinterrhein catchment (Table 1).

## 2.2 Data and data preparation

The application of bias correction algorithms to climate model outputs is generally based on three datasets: historical observations as reference (also called "target") data, historical climate model simulations, and the corresponding climate model projections. In the present study the historical reference data for the study catchments were derived from an observation-based interpolation product, i.e., the 1 km × 1 km gridded daily air temperature and precipitation datasets from the HYRAS product (Rauthe et al., 2013; Frick et al., 2014). Area-weighted mean values of precipitation and air temperature were extracted for the study catchments. The extracted catchment mean precipitation time series were corrected for undercatch based on the method by Sevruk (1989) and were then further adjusted through validation with long-term annual mean precipitation sums resulting from a water balance approach (for details see Stahl et al., 2017). The resulting time series of catchment mean precipitation and air temperature were used as input for the calibration of the glacio-hydrological model and as historically observed climate data (HOCD) for the bias correction.

The climate model datasets were obtained from the Coordinated Regional Climate Downscaling Experiment (CORDEX, http://www.cordex.org/, last access: 4 March 2019) via the Earth System Grid Federation (ESGF) archive (http://www.cordex.org/data-access/esgf/, last access: 4 March 2019). CORDEX is a collaborative effort within the climate modelling community where general circulation models (GCMs) are downscaled using regional

**Table 2.** GCM–RCM combinations from the EURO-CORDEX initiative used in this study.

| Driving GCM | RCM | RCM institution |
|---|---|---|
| CNRM-CM5-LR[1] | CCLM4-8-17 | Climate Limited-area Modelling Community |
| CNRM-CM5[1] | RCA4 | Swedish Meteorological and Hydrological Institute |
| EC-EARTH[2] | CCLM4-8-17 | Climate Limited-area Modelling Community |
| EC-EARTH[2] | HIRHAM5[5] | Danish Meteorological Institute |
| EC-EARTH[2] | RACMO22E[5] | Royal Netherlands Meteorological Institute |
| EC-EARTH[2] | RCA4 | Swedish Meteorological and Hydrological Institute |
| IPSL-CM5A-MR[3] | WRF331F | Laboratoire des Sciences du Climat et de l'Environnement |
| IPSL-CM5A-MR[3] | RCA4 | Swedish Meteorological and Hydrological Institute |
| MPI-ESM-LR[4] | CCLM4-8-17 | Climate Limited-area Modelling Community |
| MPI-ESM-LR[4] | RCA4 | Swedish Meteorological & Hydrological Institute |

GCM institutions: [1] CNRM-CERFACS (Centre National de Recherches Météorologiques-Centre Européen de Recherche et de Formation Avancée en Calcul Scientifique); note that a warning concerning an inconsistency in the historical run of CNRM-CM5 has been issued on the CORDEX errata page (https://www.hzg.de/ms/euro-cordex/078730/index.php.en, last access: 4 March 2019) after data had been downloaded and selected for this study. [2] EC-Earth consortium. [3] IPSL (Institut Pierre-Simon Laplace). [4] MPI-M (Max Planck Institute for Meteorology); [5] CORDEX errata page (https://www.hzg.de/ms/euro-cordex/078730/index.php.en, last access: 4 March 2019) notes snow accumulation issues for these GCM–RCM runs.

climate models (RCMs). Since all catchments in this study are located in Switzerland, GCM–RCM simulations were selected from the European domain of the CORDEX project (EURO-CORDEX, http://www.euro-cordex.net/, last access: 4 March 2019). EURO-CORDEX provides simulations at 0.11° (∼ 12.5 km horizontal resolution) and 0.44° (∼ 50 km horizontal resolution). Given that the catchments used in this study are situated in the Alpine domain, only the higher-resolution 0.11° simulations were used. Two Representative Concentration Pathways (RCPs) were selected for this study: RCP4.5 represents an intermediate mitigation scenario, where greenhouse gas (GHG) emissions will peak around 2040 and then steadily decrease, and RCP8.5 represents a more pessimistic scenario, which assumes that GHG emissions will continue to increase throughout the 21st century (Meinshausen et al., 2011).

Precipitation ($P$) and air temperature ($T_a$) data were provided by the 10 GCM–RCM combinations shown in Table 2 for the time period 1970–2099. For each catchment, raw GCM–RCM data were extracted using an area-weighted method as shown in Hakala et al. (2018). Based on the areal fraction of an RCM grid cell overlying a particular catchment, five RCM grid cells contribute to each catchment. All GCM–RCM combinations used in this study utilize a Gregorian calendar.

The application of the hydrological model requires catchment mean time series of $P$ and $T_a$. These were subjected to bias correction. Further data used as model input and for model calibration were not directly bias corrected. Daily potential evapotranspiration was calculated with an air-temperature-based approach provided by Oudin et al. (2005). Catchment-specific air temperature lapse rates were determined based on daily values from the HYRAS product. Based on the reference period from 1976–2006 a mean for each day of the year was calculated and smoothed using an 11-day moving average. A mean precipitation gradient (in percentage per 100 m a.s.l.) was determined from the corrected HYRAS data and applied as constant values in all simulations.

Daily streamflow data for model calibration were provided by the Swiss Federal Office for the Environment (FOEN) and the "Amt für Wasser und Abfall des Kantons Bern". The available streamflow record for the station Gündlischwand (operated by the Cantone of Berne) at the outlet of the Schwarze Lütschine study catchment covered only the period 1992–1999. By using the record of a downstream station of the Lütschine River (station Gsteig) and subtracting the streamflow of its other major headwater tributary (record from the station Zweilütschinen of the Weisse Lütschine) the streamflow for the Schwarze Lütschine study catchment could be reconstructed for the entire simulation period. This reconstructed streamflow time series was validated with the available streamflow data from the station Gündlischwand for the sub-period 1992–1999 and then used for model calibration. Snow water equivalent (SWE) and snow cover data were derived from a snow map (interpolated grid) product by the OSHD-SLF (2013). The glacier area was assessed based on glacier inventory data by Müller et al. (1976) and Maisch et al. (2000) for the state in the year 1973, by Paul et al. (2011) for the state in 2003, and by Fischer et al. (2014) for the year 2010 (see Table 1). Estimates of glacier volume were derived based on gridded ice thickness data available for the years 1973 and 2010, which were computed using the approach by Huss and Farinotti (2012) and provided by Matthias Huss. Glacier volume for the year 2003 was estimated based on the glacier cover according to Paul et al. (2011) and glacier volume–area scaling. The glacier volume estimate for 1973 was used for model initialization.

The estimate for 2003 was incorporated in the model calibration for the period 1976–2006. The estimate for 2010 was not directly used in the calibration but served the validation of model simulations beyond the year 2006.

## 3 Methods

### 3.1 Bias correction of climate data

Depending on the GCM–RCM combination, raw climate variables (noBC) of the control period (1976–2006) differ from the reference data (HOCD). To correct these biases, two different bias correction methods were applied to each climate model's $T_a$ and $P$ series: a univariate quantile mapping technique – Quantile Delta Mapping (QDM) – and a multivariate bias correction approach (MBCn). Quantile mapping is based on a transfer function that transforms the cumulative distribution(s) of the modelled data to match the distribution(s) of the observed series. The obtained transfer function is then applied to all climate model data, historical and projected. Thus it corrects systematic distributional biases relative to historical observations and preserves model-projected relative changes. QDM is a variant of quantile mapping by Cannon et al. (2015) that was designed to avoid artificial deterioration of trends arising as a statistical artefact of standard quantile mapping. QDM corrects systematic distributional biases relative to historical observations and preserves model-projected changes in quantiles in the projection period. For a given time slice, the climate model's change signal ($\Delta$) – relative change for precipitation and absolute change for air temperature – is removed from all projected future quantiles in a first step. Quantile mapping is then applied before the projected changes in quantiles are reintroduced to the bias-corrected model output.

The MBCn algorithm by Cannon (2018a) is based on the $N$-dimensional probability density function transform. This approach was originally developed for image processing (Pitié et al., 2007) but has been converted for post-processing climate model data. MBCn combines QDM and random orthogonal rotations to match the multivariate distributions of climate model data and observed data. In the MBCn approach, a random orthogonal rotation of the data points is applied before QDM. This exposes QDM to a linear combination of the original variables, which is then used to correct the marginal distributions of the rotated data. The QDM-corrected dataset is then rotated back and convergence to the observed multivariate distribution is checked. These steps are conducted iteratively until the multivariate distributions of bias-corrected climate model data and observed climate data match. In this study, 100 iterations were conducted. Both QDM and MBCn were applied in a seasonally dependent fashion. Specifically, bias corrections were applied over 30-year sliding windows. This involved replacing the central 10 years and sliding forward 10 years for each 30-year window, until the end of the projection period was reached. Within each window – to ensure an unbiased seasonal cycle – bias corrections were applied separately for each calendar month. The combination of change-preservation by QDM, which is also a core component of MBCn, with sliding windows ensures that projected trends from the underlying climate model are largely preserved. This follows the general approach and recommendation of Hempel et al. (2013) concerning trend preservation of post-processed climate model output for impact modelling.

Climate model data is often simultaneously bias corrected and downscaled as the reference data stems from stations or higher-resolution observations in comparison to the coarse grid resolution of RCMs. Undesirable effects in downscaling to finer scales have been one of the major limitations of current bias correction methods (Maraun, 2013; Ehret et al., 2012; Maraun et al., 2017). Such artefacts can occur in complex terrain in particular and if the scale gap between climate model outputs and impact model data is considerable. In general, bias correction based on spatial resolutions that differ substantially should be avoided or handled with great care. In this study the discrepancy in resolution is assumed to be acceptable as the bias correction was based on spatially aggregated mean climate variables for the meso-scale catchments (54 and 180 $km^2$) with the original resolution of the underlying gridded datasets (GCM–RCM data: 0.11°, historical HYRAS data: 1 km) becoming of secondary importance.

### 3.2 Hydrological model simulations

The HBV model (Bergström, 1976; Lindström et al., 1997) is a semi-distributed bucket-type runoff model. Here the software implementation HBV-light (Seibert and Vis, 2012) was used, which recently has been extended to represent coupled glacio-hydrological processes of partly glacierized catchments (Seibert et al., 2018). This version of the HBV model also allows tracking of the different components of streamflow resulting from rainfall ($Q_R$), snowmelt ($Q_S$), and glacier ice melt ($Q_I$) (Weiler et al., 2018; Seibert et al., 2018). The HBV model requires daily precipitation, air temperature, and potential evapotranspiration data as input to simulate daily runoff. In addition, linear gradients of air temperature and precipitation are needed for the interpolation over elevation zones. A general description of the basic model structure and the process conceptualization of the HBV model are found elsewhere (e.g., Lindström et al., 1997; Seibert and Vis, 2012; Seibert et al., 2018). Snow and ice accumulation and melt are based on a widely used air temperature index approach using a threshold air temperature as a model parameter to differentiate between precipitation falling as snow and rain as well as to simulate melt of snow and ice by additionally using a degree-day factor. Differences in the melt of glacier ice compared to snow are represented by another model parameter. The influence of differences in aspect on snow and ice melt was taken into account by distinguishing

**Table 3.** Model performance criteria for the calibration (1 October 1976–30 September 2003) and validation (1 October 2003–31 December 2006) of the hydrological model formulated (see footers) that the ideal value for a perfect fit is 1.0.

| Model performance criteria | Weight in | Hinterrhein | | Schwarze Lütschine | |
|---|---|---|---|---|---|
| | calibration | Calibration | Validation | Calibration | Validation |
| Nash–Sutcliffe efficiency ($R_{\text{eff}}$)[1] for streamflow | – | 0.773 | 0.763 | 0.910 | 0.880 |
| Kling–Gupta efficiency[2] for streamflow | – | 0.861 | 0.877 | 0.934 | 0.898 |
| Volume error ($V$)[3] for streamflow | – | 0.972 | 0.962 | 1.000 | 0.965 |
| Lindström measure[4] for streamflow | 0.20 | 0.770 | 0.759 | 0.910 | 0.877 |
| $R_{\text{eff}}$[1] for log transformed streamflow | 0.15 | 0.840 | 0.648 | 0.908 | 0.749 |
| $R_{\text{eff}}$[1] for streamflow in Jun–Sep | 0.15 | 0.684 | 0.711 | 0.795 | 0.749 |
| Root mean square error for snow-covered area fraction[5] | 0.10 | 0.856 | 0.761 | 0.863 | 0.803 |
| Mean absolute normalized error (MANE) for SWE[6] | 0.20 | 0.642 | 0.557 | 0.757 | 0.553 |
| Glacier volume change objective function[7] | 0.20 | 0.999998 | – | 0.999994 | – |

Formulation of model performance criteria: [1] $R_{\text{eff}} = 1 - \frac{\sum(Q_{\text{obs}} - Q_{\text{sim}})^2}{(Q_{\text{obs}} - \overline{Q_{\text{obs}}})^2}$ where $Q_{\text{obs}}$ and $Q_{\text{sim}}$, respectively, are observed and simulated streamflow (mm day$^{-1}$). [2] See Gupta et al. (2009). [3] $1 - V$ with $V = \frac{\sum|(Q_{\text{obs}} - Q_{\text{sim}})|}{\sum(Q_{\text{obs}})}$ where $Q_{\text{obs}}$ and $Q_{\text{sim}}$, respectively, are observed and simulated streamflow (mm day$^{-1}$). [4] $1 - R_{\text{eff}} - 0.1V$ with $R_{\text{eff}}$[1] and $V$[6]; see Lindström et al. (1997). [5] $1 - \sqrt{\frac{1}{n}(C_{\text{sim}} - C_{\text{ref}})^2}$ with $C_{\text{ref}}$ the snow-covered catchment area fraction ($C$) (–) as per gridded SWE reference data, $C_{\text{sim}}$ the simulated $C$, and $n$ the number of time steps. [6] $1 - \frac{\sum|(S_{\text{ref}} - S_{\text{sim}})|}{\sum S_{\text{obs}}}$ with $S$ (mm) the mean SWE for elevation range below 2500 m a.s.l. where $S_{\text{ref}}$ is derived from SWE reference data and $S_{\text{sim}}$ is simulated. [7] $1 - \frac{|\Delta W_{\text{sim}} - \Delta W_{\text{obs}}|}{\Delta W_{\text{obs}}}$ with $\Delta W$ (mm) the change of glacier ice volume in water equivalent between the years 1973 and 2003, where $\Delta W_{\text{obs}}$ corresponds to an estimate based on observed glacier area and $\Delta W_{\text{sim}}$ is simulated.

three aspect classes and applying an additional aspect factor parameter (Hagg et al., 2007; Hottelet et al., 1993). The latest version of the HBV-light software with the implementation of the coupled glacio-hydrological processes and the adjustment of glacier geometry to glacier mass changes based on the $\Delta h$ parametrization by Huss et al. (2010) is explained in detail in Seibert et al. (2018). It should be noted that with the implementation in HBV-light only one glacier per catchment or subcatchment can be represented. Hence, glacier cover areas in each of the two case study catchments were aggregated and simulated as one "virtual" model glacier.

The model was calibrated for the period from 1976–2003, preceded by a 3-year warm-up period, by optimizing a weighted objective function, giving special attention to streamflow dynamics (50 %), snow simulation (25 %), and glacier volume change (25 %). The Lindström measure (Lindström, 1997) was used for the streamflow's general dynamic and volume errors, while the Nash–Sutcliffe efficiency (Nash and Sutcliffe, 1970) was computed based on logarithmically transformed streamflow. Additionally the Nash–Sutcliffe efficiency was computed for the streamflow only during the summer months, from June to September. To calibrate the snow simulations the snow-covered area fraction of the catchment and the mean SWE of the elevation range < 2500 m a.s.l. were used. Elevations below 2500 m a.s.l. represent the crucial range for the snow line and in this range the gridded SWE interpolation used as reference data is well-founded on station data. Glacier volume was considered in the calibration process using glacier volume estimates for the years 1973 and 2003. The automated multi-criteria calibration was based on a genetic algorithm for parameter optimization (see Seibert, 2000). A 3-year model validation period (1 October 2003–31 December 2006) completed the historical reference period 1977–2006. Resulting performance measures for the calibration and validation period are summarized in Table 3 (see Supplement for additional figures comparing simulated variables and reference data). The retreat of the glaciers required all experiments to be run in a transient mode, i.e., the model was forced with climate model scenario data for the period from October 1976 to September 2099.

### 3.3 Data analysis

Effects of the bias correction approaches on the hydrological simulation were based on comparisons of the simulation results for the historical reference period 1976–2006 using $P$ and $T_{\text{a}}$ time series derived from the HYRAS datasets as input ($\text{Sim}_{\text{HOCD}}$) and simulations forced with $P$ and $T_{\text{a}}$ series from 10 different GCM–RCM outputs for the two different RCP scenarios, each uncorrected ($\text{Sim}_{\text{noBC}}$) and bias corrected based on QDM ($\text{Sim}_{\text{QDM}}$) and on MBCn ($\text{Sim}_{\text{MBCn}}$). In total, this led to 61 hydrological model runs (1 $\text{Sim}_{\text{HOCD}}$, 20 $\text{Sim}_{\text{noBC}}$, 20 $\text{Sim}_{\text{QDM}}$, and 20 $\text{Sim}_{\text{MBCn}}$) per catchment. In a first step (Sect. 4.1), the different $P$ and $T_{\text{a}}$ series were evaluated for the amount of precipitation occurring at air temperatures above and below 0° C due to the importance for the simulation of snow accumulation and melt processes. Furthermore, the simulation results were assessed in terms of SWE, glacier ice volume ($V_{\text{I}}$) evolution (Sect. 4.2), and even-

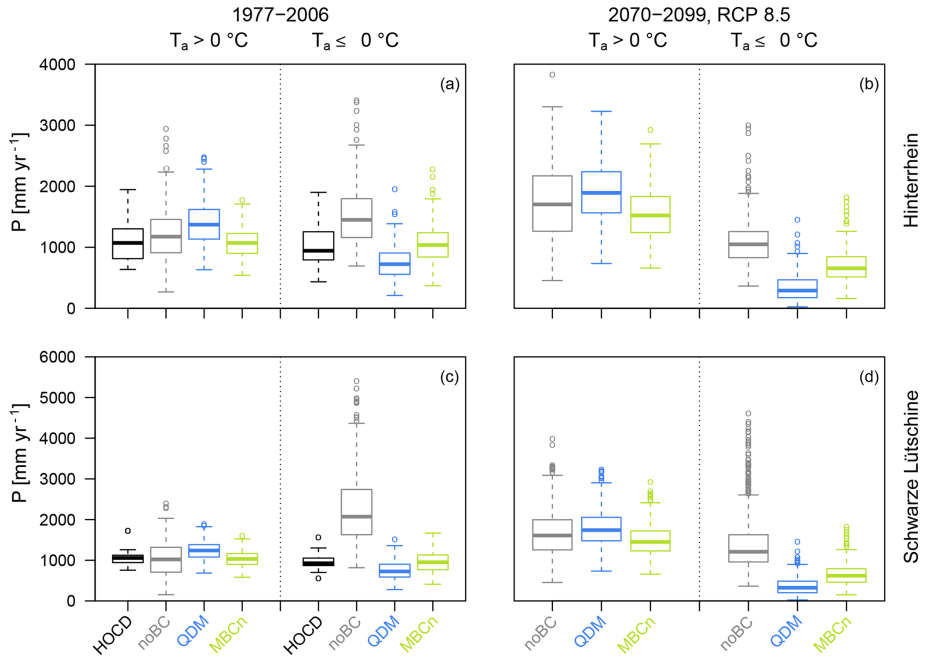



**Figure 2.** Annual precipitation sums for days with air temperatures above or below 0 °C.

tually streamflow with its three individual components $Q_R$, $Q_S$, and $Q_I$ (Sect. 4.3).

## 4 Results

### 4.1 Climate variables bias correction

The two applied bias correction methods led to differences concerning the interdependence of $P$ and $T_a$. The distribution of annual precipitation sums during air temperatures above and below 0 °C in the entire ensemble is represented in Fig. 2, while results for the individual GCM–RCM output series are provided in the Supplement. Generally, the uncorrected climate model data (noBC) have a wider variability than the reference data (HOCD). Particularly for the Schwarze Lütschine the uncorrected data yielded precipitation amounts remarkably higher than historically observed. However, differences also existed between the correction methods. For both catchments precipitation falling above air temperatures of 0 °C was overestimated with QDM. Accordingly, precipitation falling below air temperatures of 0 °C was underestimated in the univariate bias-corrected data. MBCn appears to have better reproduced the historical reference data in this respect.

### 4.2 Hydrological model simulations – cryosphere

Application of the climate scenarios clearly revealed a decreasing role of snow for both study catchments. Figure 3 illustrates a distinctly smaller snow accumulation in the course of a year simulated for the period 2070–2099 (compared to the historical reference period 1977–2006) and a more complete melt during the summer. This extended the snow-free period during the summer in the Hinterrhein catchment. The spread between the simulations diverged for the simulations of future conditions. In the Schwarze Lütschine catchment with its higher maximum elevations all effects were comparable, yet a permanent snow cover remained still present based on most scenarios. As expected, simulations based on the RCP4.5 scenario (not shown) led to a clear but less severe decrease in mean SWE than for the RCP8.5 scenario.

The differences in the interdependence of precipitation and air temperature resulting from the application of QDM versus MBCn to the GCM–RCM data can be seen in the simulated SWE (Fig. 3). The state of precipitation defined by the calibrated threshold air temperature parameter TT (Schwarze Lütschine TT = −0.29 °C; Hinterrhein TT = −0.73 °C) influenced the snow accumulation and therefore led to differences in the annual SWE regime (Fig. 3). As MBCn-corrected GCM–RCM data caused more precipitation to fall as snow, the accumulated catchment mean SWE in spring was simulated to be up to around 100–200 mm higher in the historical reference period compared to simulations based on QDM-corrected forcing data. Simulated SWE based on the two different bias correction methods differed notably. Comparing the results with the reference simulation (Fig. 3) indicates that MBCn performed better. The systematic difference in simulated SWE resulting from the bias correction methods was a bit less clear for the Schwarze Lütschine catchment in the scenario period, yet overall the differing tendencies between QDM- and MBCn-corrected data were considerable.

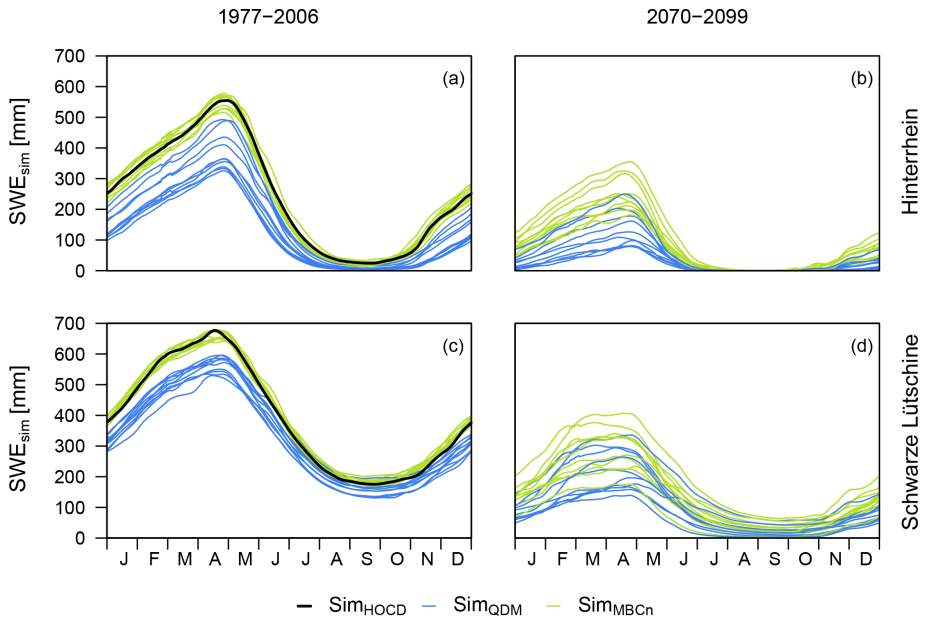



**Figure 3.** Mean annual SWE regime, calculated using the 11-day moving average of daily simulated SWE (catchment mean) during the historical reference period **(a** and **c)** and at the end of the scenario period based on the RCP8.5 scenario **(b** and **d)**.

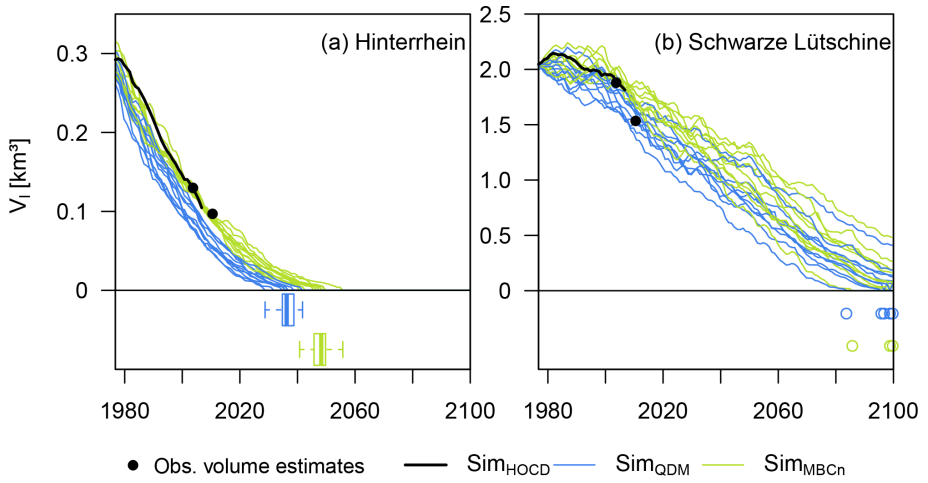



**Figure 4.** Simulated glacier ice volume from 1977 to 2099 using the RCP8.5 scenario forcing in the two catchments **(a, b)**. In the lower part of the graphs the boxes in the left figure and the dots in both figures indicate the simulated years of the complete glacier ice melt. For the Schwarze Lütschine only 5 (3) out of the 10 Sim$_{QDM}$ (Sim$_{MBCn}$) simulations led to complete glacier melt by the end of 2099, not allowing any box plots to be shown. Filled black circles are glacier volume estimates based on observed glacier area data in 2003 and 2010.

For the period 1976 to 2099 the glacier volume was simulated to decrease in both catchments. In the Hinterrhein catchment, glaciers diminished continuously from the beginning of the simulation period and were simulated to have disappeared between 2028 and 2055 under the RCP 8.5 scenario depending on the GCM–RCMs and the applied bias correction method (Fig. 4). In the Schwarze Lütschine catchment, data from a few GCM–RCMs resulted in an increase in simulated glacier volume in the 1970s and 1980s, which is in line with the historical reference simulation (Sim$_{HOCD}$). In the following years, glacier volume decreased continuously. In contrast to the Hinterrhein catchment, glaciers were not simulated to have disappeared by the end of 2099 based on the RCP4.5 scenario (not shown). However, in the simulations the glacier volume diminished to on average roughly a third of its initial size at the beginning of the simulation period. The RCP8.5 scenario from a few certain GCM–RCM combinations even led to complete glacier disappearance in the Schwarze Lütschine catchment within the 21st century.

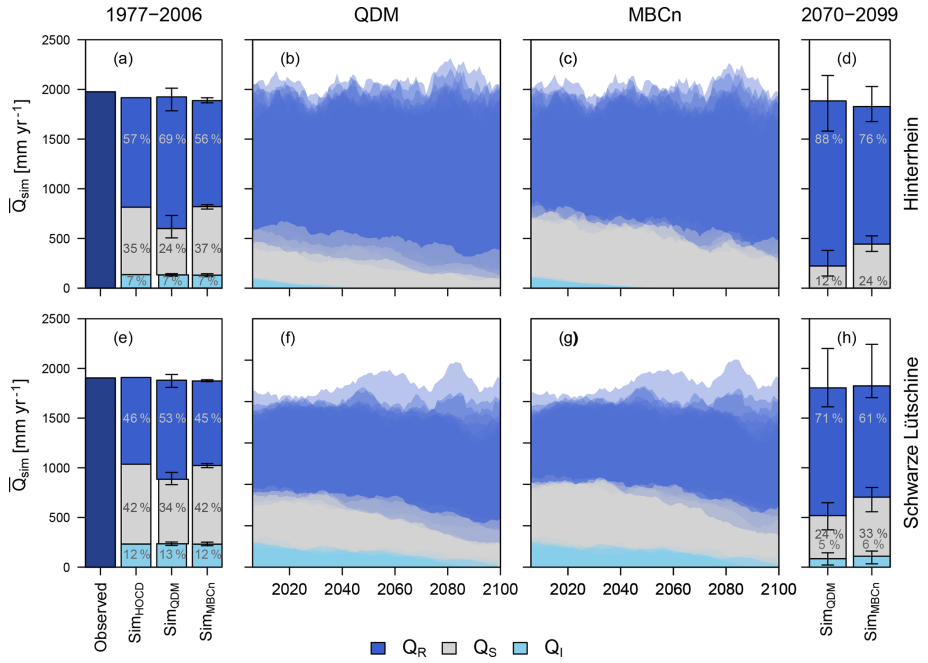



**Figure 5.** Observed total streamflow and simulated streamflow components for the historical reference period and for the different simulations under the RCP8.5 scenario. Stacked bar plots show mean values over the historical reference period **(a, e)** and for the period 2070–2099 **(d, h)**, stacked bar plots for $Sim_{QDM}$ and $Sim_{MBCn}$ show ensemble mean with ensemble spread (error bars). Simulation results over the scenario period 2006–2099 **(b, c, f, g)** are shown as semi-transparent polygons for each GCM–RCM combination.

Focusing on systematic differences between simulations using data corrected based on QDM and MBCn, the simulations of glacier volume showed similar tendencies to those found for SWE. For both catchments, but again more clearly for the Hinterrhein catchment, MBCn-corrected GCM–RCM data resulted in a slower decline in glacier volume in comparison to simulations based on QDM-corrected data. All projections led to complete glacier disappearance in the Hinterrhein catchment by about the year 2050, with a clear tendency towards earlier dates for QDM-based simulations (2028–2041, mean: 2036) compared to MBCn-based simulations (2040–2055, mean: 2047). For the Schwarze Lütschine catchment the range of QDM- and MBCn-based glacier volume simulations overlapped largely as simulations in general diverged considerably. However, for each individual GCM–RCM dataset, glacier melt was simulated to be faster using the QDM-corrected data compared to the MBCn-corrected data. The less intense decline in glacier volumes resulting from MBCn-corrected forcing data appeared to correspond better with the reference simulation ($Sim_{HOCD}$) in the initial phase of the historical period and with the observation-based glacier volume estimates for the year 2003 (and also for the year 2010 in case of the Hinterrhein catchment). MBCn thus led to more realistic results for the historical reference period.

### 4.3 Hydrological model simulations – streamflow

Time changes of annual variables and mean monthly hydrological regimes were assessed for streamflow $Q$ and for the individual streamflow components, i.e., the rain component $Q_R$, the snowmelt component $Q_S$, and the ice melt component $Q_I$. Mean annual streamflow of the study catchments showed a small decrease over the entire simulation period from 1976 to 2099 for most simulations, while for some a slight increase was noticed (Fig. 5). However, the simulations based on different GCM–RCM outputs diverged over time. While, on average, the total annual streamflow stayed largely unchanged, its composition clearly changed. The streamflow component from glacier ice melt decreased slowly over time as the glaciers retreated. Likewise, the snowmelt component of streamflow decreased over time. On average, for the RCP4.5 scenario's MBCn-corrected data these decreases were around 14 % in the Hinterrhein and 16 % in the Schwarze Lütschine for the RCP8.5 scenario's QDM-corrected data they were around 53 % in the Hinterrhein and 33 % in the Schwarze Lütschine.

The streamflow simulations reflected the changes from the different bias correction methods found for the cryosphere. Simulations based on QDM-corrected data led to slightly different total streamflow than MBCn-corrected data (Fig. 5a, d and e). These differences were much more pronounced regarding the individual streamflow components. Modelling based on QDM-corrected climate data led to an approxi-

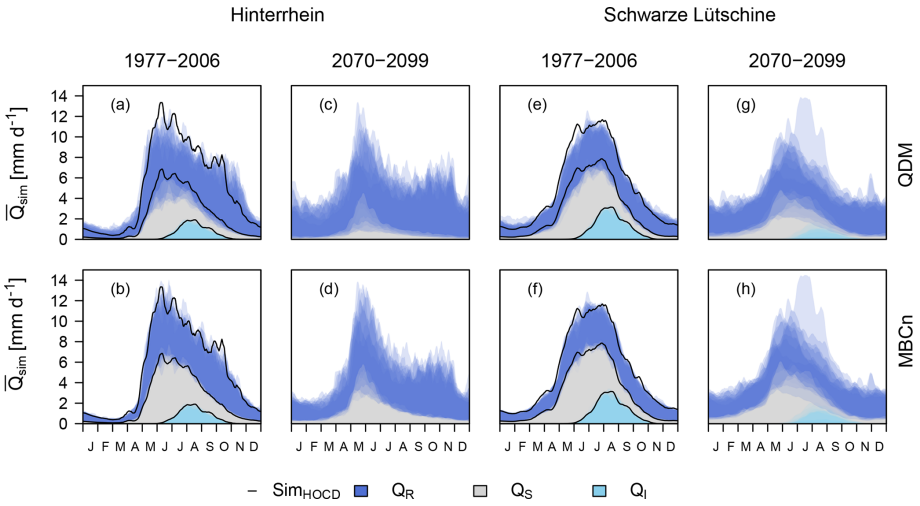



**Figure 6.** Streamflow regimes based on 11-day moving averages of daily streamflow during 30-year periods in the historical reference period and as projected for the period 2070–2099 under the RCP8.5 scenario for the two catchments. Simulation results for each ensemble member are shown as semi-transparent polygons. For the historical reference period the results of the simulations based on the historical reference $P$ and $T_a$ time series are also shown (black lines).

mately 10 % higher rain component of streamflow $Q_R$ in comparison to MBCn-corrected simulations. The snowmelt component of streamflow $Q_S$ varies proportionally, being notably smaller when using QDM-corrected GCM–RCM data. Comparing the means of the ice melt components of streamflow $Q_I$ for the 30-year periods at the beginning and at the end of the entire simulation period showed no differences from the bias correction methods for the Hinterrhein catchment and differences in the range of only 1 % for the Schwarze Lütschine catchment.

Simulated streamflow and its components, $Q_I$, $Q_S$, and $Q_R$, also changed seasonally (Fig. 6). In the historical reference period (1977–2006), the two catchments had a nivo-glacial streamflow regime peaking in the summer due to snow and ice melt and with little streamflow during winter. According to the projections the streamflow peak in early summer remained a dominant characteristic until the end of the simulation period. Yet, for the Hinterrhein catchment, the peak's timing was simulated to shift, causing streamflow to concentrate in May and the peak to become much narrower than in the past. For the Schwarze Lütschine catchment the simulations for the RCP8.5 scenario resulted in very variable summer streamflow regimes for 2070–2099 and a tendency towards a lower summer streamflow peak than in the past. In the reference period, the glaciers' influence showed during late summer, where it extended the melt peak into autumn. This effect was simulated to diminish, resulting in decreased total streamflow in late summer. During autumn and winter, simulated streamflow for 2070–2099 was nearly double the level of the historical period mainly due to an increase in the rainfall component of streamflow. Despite similar tendencies of reduced $Q_S$ in the future, differences arising from the different bias correction methods are notable. $Q_S$ was more prominent in all regimes based on MBCn-corrected GCM–RCM outputs, which simulated higher peaks during the snowmelt season and a generally higher fraction during the rest of the year, especially for the future periods. Accordingly, QDM-corrected data supported a larger $Q_R$ component beyond the summer. As a consequence, during low-flow periods in winter, QDM-corrected forcing data overestimated the streamflow in the historical reference period. In contrast, QDM-corrected forced simulations tended to slightly underestimate the streamflow during the spring and summer months, as $Q_S$ was underestimated. Generally, MBCn-corrected data matched more closely with the reference simulations based on observed data.

## 5 Discussion

Both bias correction methods employed in this study, univariate QDM (Cannon et al., 2015) and multivariate MBCn (Cannon, 2018a), are based on the same quantile mapping approach and by definition the marginal distributions of the corrected $P$ and $T_a$ series are the same as those of the historical reference data in the historical period; furthermore, the two methods also share the same marginal distributions in the projection period. However, the bias correction methods do result in differences in terms of $P$ and $T_a$ interdependency (see marginal and joint distributions of $P$ and $T_a$ series in the Supplement). Preserving the ranks of the climate model simulations, univariate bias correction approaches retain the inter-variable dependencies as represented in the raw climate model output (Vrac, 2018), as also demonstrated and discussed in previous studies using univariate quantile mapping methods (Wilcke et al., 2013; Ivanov and Kot-

larski, 2017). However, observed inter-variable dependencies are often misrepresented in climate model simulations and hence biases therein are also retained by univariate methods (Wilcke et al., 2013; Gennaretti et al., 2015; Zscheischler et al., 2019). Such biases in the interdependency representation were also found in this study for $P$–$T_a$ interdependency of raw climate model output from 10 GCM–RCM simulations compared to the historical observational dataset used (see also Supplement). In snow-dominated environments, the representation of precipitation–temperature interdependence is important for hydrological modelling but also for many other aspects impacted strongly by snow cover extent and duration (Gennaretti et al., 2015). Further studies that compare $P$–$T_a$ representation in climate model output and multiple observational datasets are needed to explore the causes of differences between climate model output and reference data such as those found here.

As air temperature determines the distinction between liquid precipitation and snow, differences in the climate variables' interdependence can lead to differences in simulated snowfall (Fig. 2), and consequently in snow accumulation and the catchments' seasonal water storage (Figs. 3–6). For the MBCn-corrected data in this study there was clearly more precipitation at air temperatures below 0 °C in comparison to the QDM-corrected data, resulting in more precipitation falling as snow and being stored and accumulated than for univariate bias-corrected forcing data. In glacierized catchments the higher amounts of snow from MBCn compared to QDM also affected the glaciers, with higher winter mass balances and a later start of the melt season in spring and summer. The existence or non-existence of water storages in the form of snow and ice as well as the liquid precipitation directly contributing to streamflow had notable influences on the streamflow composition and regime. For instance, the larger fraction of liquid precipitation at the cost of snow simulated with QDM-corrected data led to a systematic overestimation of streamflow during the winter months in the historical reference period. This error was not present in simulations based on MBCn-corrected $P$ and $T_a$ forcing.

It bears noting that results from QDM and MBCn in the historical reference period are, as for example also in Zscheischler et al. (2019), evaluated without cross-validation. However, because the univariate and multivariate bias correction algorithms are applied in an asynchronous fashion to freely running climate simulations – adjusting the marginal and joint distributions – it is, by construction, almost guaranteed that they will perform well in terms of cross-validated measures of distributional fit (Maraun and Widmann, 2018). Cross-validation does make sense when performance – especially for aspects not explicitly adjusted – is measured in a setting where climate model simulations are synchronized with the real-world climate state, for example in climate prediction or perfect boundary condition (e.g., reanalysis-driven) setups. We note that such reanalysis-driven cross-validation experiments have been performed in Cannon (2018a) for the two algorithms used in this study. This was done over a large continental domain for a complicated multivariate fire weather index that combines, in a non-linear fashion, the current and lagged effects of air temperature, precipitation, wind, and humidity. Hence, it is expected that results reported here are robust and would be similar in an out-of-sample evaluation.

There have long been concerns over climate change impacts on mountain water towers. Many climate impact studies for snow-dominated catchments agree that due to continued warming, a decrease in snow cover characteristics and time-shifted snowmelt contributions to streamflow are to be expected under climate change scenarios (e.g., Barnett et al., 2005; Farinotti et al., 2012; Köplin et al., 2014; Addor et al., 2014; Milano et al., 2015; Coppola et al., 2018; Jenicek et al., 2018; Hanzer et al., 2018). In fact, the shift and decrease in the snowmelt peak are one of the most robust results of such studies. In this study we showed that the snow component strongly depends not only on the GCM–RCM outputs but also on whether the bias correction method applied incorporates inter-variable dependence of $P$ and $T_a$ or not. The simulated glacier volume showed a clearly decreasing trend over the scenario period. However, net mass balances and hence rates of glacier ice melt and the mean timing of the final glacier disappearance vary by over a decade in the Hinterrhein catchment. While the ensemble covers a wide range, the bias correction approach makes a difference for each GCM–RCM forcing. The changes in snow accumulation and glacier melt then propagate into changes of streamflow regimes. In future projections, snowmelt peaks tend to occur earlier and with a more concentrated melt season. A potential effect of this storage shift on streamflow, however, is potentially relevant year-round, as could be visualized by the specific streamflow component modelling. The simulations suggest that the melt contribution to streamflow depends on the interdependence of air temperature and precipitation and hence the chosen bias correction method. Furthermore, streamflow during the late summer decreases as the release of stored water from glaciers, which makes up a notable percentage of streamflow during the late summer, will have diminished. These systematic differences in hydrological impact scenarios originating from the applied univariate or multivariate bias correction method such as those found here, e.g., differences in glacier disappearance dates or differences in seasonal (summer vs. winter) water availability, may appear negligible given the overall large uncertainties of climate impact modelling and yet may still be relevant for some specific adaptation management questions. The timing of "peak water" occurrence or complete disappearance of glaciers may be relevant for the planning horizon of hydropower schemes (Hänggi and Weingartner, 2012; Schaefli, et al., 2019). The earlier recession of the melt peak may sooner or later affect early-summer flood hazards or increase the hazard of late-summer low flows due to the loss of ice and snow components of streamflow (Beaulieu et al., 2012; God-

sey et al., 2014), requiring the planning of respective measures.

These results also require the discussion of implications on common conceptual hydrological modelling concepts that are needed to simplify meteorological and hydrological complexity. The use of a threshold air temperature for the distinction of precipitation in snow and rainfall is a key concept of the HBV model and many other hydrological models. Hence, it may be expected that the simulations of the snow-dominated catchments respond particularly sensitively to changes and biases in $P$–$T_a$ interdependencies. The question is the degree to which this may influence the hydrological variables discussed above. So far, few studies have evaluated multivariate-corrected GCM–RCM data in hydrological modelling. Chen et al. (2018) found that the joint bias correction of precipitation and air temperature led to a much better performance in terms of hydrological modelling for all their study basins located in various climates except for the coldest Canadian basin. In contrast, an overall additional benefit of using bivariate bias correction methods for hydrological impact projections was not evident in results by Räty et al. (2018) when compared to using a univariate quantile mapping applied as a delta change method, i.e., retaining present-day correlation structures. However, their analysis indicated that the selection of the bias correction method was most important and the added value of using multivariate approaches most clearly found for SWE simulations, supporting the findings of this study. Based on these case studies, it may be assumed that simulations with any hydrological model that include calibration over a historical reference period will be somewhat affected by a biased representation of the inter-variable dependence of its input variables in GCM–RCM outputs. Further studies are needed to investigate other effects of multivariate bias correction for other types of climatological input variables, hydrological models, catchment types, and dominating processes.

This study demonstrates the importance of considering the representation of the interdependence of precipitation and air temperature in the specific case of hydrological impact modelling of snow- and glacier-dominated catchments. As shown, in the representation of the climate variables' interdependence, the multivariate bias correction approach leads to results closer to the climatological historical reference data as well as partly to hydrological simulations closer to the historical reference simulations, such as for instance for the simulated glacier volumes. Cannon (2016, 2018a) also demonstrated better results for multivariate-corrected data in other examples, including fire weather indices and atmospheric river detection. In practice, some kind of bias correction is needed for many impact studies, although it is known that recent literature is rich in controversial debate of its use and major limitations of the application of empirical–statistical bias correction methods (e.g., Ehret et al., 2012; Addor and Seibert, 2014; Maraun, 2013, 2016; Clark et al., 2016; Maraun et al., 2017; Casanueva et al., 2018; Zscheischler et al., 2019). Some of the fundamental issues, the details of which are beyond the scope of this study, are shared with univariate bias correction, for example, the question of stationarity (regarding biases in marginal distributions). In addition, joint correction is often based on the assumption that the structure of the bias in variables' interdependence is stationary, i.e., the same for control as for projections. This is not strictly true for MBCn, which allows the multivariate distribution to evolve in the projection period. However, the extent to which model-projected changes in dependence structure are preserved by MBCn has yet to be evaluated closely. More generally, whether the preservation of inter-variable dependence structures is a robust assumption or dependence structures should evolve from the reference to the future period are still open questions for the development of multivariate bias correction methods (Vrac, 2018). Furthermore, the correction of the multivariate dependence structure will necessarily affect the time sequencing of the climate model variables (Cannon, 2016), which can lead to modification of temporal autocorrelation. Maraun (2016) cautions that modifications of spatial, temporal, or multi-variable interdependence may break the consistency with the driving climate model and many others have argued for the least possible transformation of GCM–RCM outputs for this reason. This study does not address these fundamental questions and critiques nor does it generally recommend or not recommend the use of multivariate bias correction methods. The objective of the study was to compare the differences resulting from univariate vs. multivariate methods. We demonstrated a case in which biases in inter-variable dependencies can affect hydrological simulations considerably. This is important, particularly as it is common practice to use hydrological models calibrated to climatic conditions represented by historical climate variable series. In the same way that the use of several climate and hydrological models is recommended, the incorporation of uncorrected and univariate- and multivariate-corrected scenario data in the ensemble may be considered as one part of a transparent and honest communication of the full range of uncertainties.

## 6 Conclusions

This study systematically tested the effects of multivariate bias correction of projected air temperature and precipitation versus a traditional univariate bias correction on hydrological impact modelling in alpine environments. Jointly corrected air temperature and precipitation series simulated more snowfall and consequently up to 50 % more snow accumulation than univariate-corrected GCM–RCM data. Subsequently, glacier volume was simulated to decrease by up to a decade slower under multivariate-corrected scenarios. These differences also impact the simulations of streamflow and its components with higher snowmelt components and accordingly smaller rainfall components under multivariate-

corrected scenarios compared to univariate-corrected scenarios. These are relevant systematic differences despite variations in the GCM–RCM ensemble. The choice between a univariate and a multivariate bias correction approach may therefore have implications for future water resources planning, as the snow component presents an important seasonal storage, and for the protection against hydrological hazards such as a higher vulnerability to drought.

Beyond this specific case this study suggests that the effect of bias correction methods may be generalized for catchments that include the elevation range of the snow line. Mountain hydrology modelling relies on the correct representation of the interdependence of air temperature and precipitation due to a crucial role of threshold air temperature concepts for the distinction of liquid and solid precipitation. This study makes an argument for the explicit consideration of interdependencies of climate variables by using multivariate bias correction methods in hydrological climate change impact studies in snow-dominated catchments. But many other threshold effects also drive relevant climate impacts and are parameterized in many models or indices. The study provides a strong incentive to test similar effects in hydrological systems and their model representations that may be dominated by other climate variable interdependencies.

*Code availability.* An R package (R Core Team, 2018) including the MBCn and the QDM algorithm is available for download from https://CRAN.R-project.org/package=MBC (Cannon, 2018b). The HBV-light software is freely available for download from https://www.geo.uzh.ch/en/units/h2k/Services/HBV-Model.html.

*Data availability.* EURO-CORDEX data can be accessed via different European data nodes, available at https://www.hzg.de/ms/euro-cordex/060378/index.php.en (last access: 4 March 2019). The HYRAS interpolation product used to derive the historical reference climate time series was made available by the German Weather Service (DWD) and the German Federal Institute of Hydrology (BfG). Streamflow time series were provided by the Swiss Federal Office for the Environment (FOEN) and the Amt für Wasser und Abfall des Kantons Bern. Snow data of the "SLF-Schneekartenserie Winter 1972–2012" used for model calibration are available upon request by the WSL Institute for Snow and Avalanche Research (SLF). Glacier ice thickness data were provided by Matthias Huss, and other glacier data are available according to the given references.

*Supplement.* The supplement related to this article is available online at: https://doi.org/10.5194/hess-23-1339-2019-supplement.

*Author contributions.* JM, IK, KS, and JS designed the study. JM carried out bias correction, modelling, and all analyses and wrote the first draft. IK calibrated the hydrological model and prepared snow, glacier, and hydrological data. KH prepared the EURO-CORDEX data for the catchments. AC provided and helped with his bias correction scripts. All co-authors contributed to and edited the paper.

*Competing interests.* The authors declare that they have no conflict of interest.

*Acknowledgements.* Work for this study was based on data acquired and methods developed within the project "The snow and glacier melt components of the streamflow of the River Rhine and its tributaries considering the influence of climate change" (ASG-Rhein, see Stahl et al., 2017) funded by the International Commission for the Hydrology of the Rhine basin (CHR). We thank Urs Beyerle for his assistance with the retrieval of EURO-CORDEX data and further thank all data providers (see Data availability). The article processing charge was funded by the German Research Foundation (DFG) and the University of Freiburg in the funding programme Open Access Publishing. Valuable comments by the editor and the reviewers helped to improve the paper.

Edited by: Luis Samaniego
Reviewed by: Ole Rössler and Sven Kotlarski

## References

Addor, N. and Seibert, J.: Bias correction for hydrological impact studies – beyond the daily perspective, Hydrol. Process., 28, 4823–4828, https://doi.org/10.1002/hyp.10238, 2014.

Addor, N., Rössler, O., Köplin, N., Huss, M., Weingartner, R., and Seibert, J.: Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments, Water Resour. Res., 50, 7541–7562, https://doi.org/10.1002/2014WR015549, 2014.

Barnett, T. P., Adam, J. C., and Lettenmaier, D. P.: Potential impacts of a warming climate on water availability in snow-dominated regions, Nature, 438, 303–309, https://doi.org/10.1038/nature04141, 2005.

Beaulieu, M., Schreier, H., and Jost, G.: A shifting hydrological regime: A field investigation of snowmelt runoff processes and their connection to summer base flow, Sunshine Coast, British Columbia, Hydrol. Process., 26, 2672–2682, https://doi.org/10.1002/hyp.9404, 2012.

Bergström, S.: Development and application of a conceptual runoff model for Scandinavian catchments, SMHI, Norrköping, 1976.

Bosshard, T., Kotlarski, S., Zappa, M., and Schär, C.: Hydrological Climate-Impact Projections for the Rhine River: GCM–RCM Uncertainty and Separate Temperature and Precipitation Effects, J. Hydrometeorol., 15, 697–713, https://doi.org/10.1175/JHM-D-12-098.1, 2014.

Cannon, A. J.: Multivariate Bias Correction of Climate Model Output: Matching Marginal Distributions and Intervariable Dependence Structure, J. Climate, 29, 7045–7064, https://doi.org/10.1175/JCLI-D-15-0679.1, 2016.

Cannon, A. J.: Multivariate quantile mapping bias correction: An $N$-dimensional probability density function transform for cli-

mate model simulations of multiple variables, Clim. Dynam., 50, 31–49, https://doi.org/10.1007/s00382-017-3580-6, 2018a.

Cannon, A. J.: Multivariate Bias Correction of Climate Model Outputs, R package 'MBC' version 0.10-4, available at: https://CRAN.R-project.org/package=MBC (last access: 4 March 2019), 2018b.

Cannon, A. J., Sobie, S. R., and Murdock, T. Q.: Bias Correction of GCM Precipitation by Quantile Mapping: How Well Do Methods Preserve Changes in Quantiles and Extremes?, J. Climate, 28, 6938–6959, https://doi.org/10.1175/JCLI-D-14-00754.1, 2015.

Casanueva, A., Bedia, J., Herrera, S., Fernández, J., and Gutiérrez, J. M.: Direct and component-wise bias correction of multi-variate climate indices: The percentile adjustment function diagnostic tool, Climatic Change, 121, 2075, https://doi.org/10.1007/s10584-018-2167-5, 2018.

Chen, J., Brissette, F. P., Chaumont, D., and Braun, M.: Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America, Water Resour. Res., 49, 4187–4205, https://doi.org/10.1002/wrcr.20331, 2013.

Chen, J., Li, C., Brissette, F. P., Chen, H., Wang, M., and Essou, G. R. C.: Impacts of correcting the inter-variable correlation of climate model outputs on hydrological modeling, J. Hydrol., 560, 326–341, https://doi.org/10.1016/j.jhydrol.2018.03.040, 2018.

Clark, M. P., Wilby, R. L., Gutmann, E. D., Vano, J. A., Gangopadhyay, S., Wood, A. W., Fowler, H. J., Prudhomme, C., Arnold, J. R., and Brekke, L. D.: Characterizing Uncertainty of the Hydrologic Impacts of Climate Change, Curr. Clim. Change Rep., 2, 55–64, https://doi.org/10.1007/s40641-016-0034-x, 2016.

Coppola, E., Raffaele, F., and Giorgi, F.: Impact of climate change on snow melt driven runoff timing over the Alpine region, Clim. Dynam., 51, 1259–1273, https://doi.org/10.1007/s00382-016-3331-0, 2018.

Ehret, U., Zehe, E., Wulfmeyer, V., Warrach-Sagi, K., and Liebert, J.: HESS Opinions "Should we apply bias correction to global and regional climate model data?", Hydrol. Earth Syst. Sci., 16, 3391–3404, https://doi.org/10.5194/hess-16-3391-2012, 2012.

Farinotti, D., Usselmann, S., Huss, M., Bauder, A., and Funk, M.: Runoff evolution in the Swiss Alps: Projections for selected high-alpine catchments based on ENSEMBLES scenarios, Hydrol. Process., 26, 1909–1924, https://doi.org/10.1002/hyp.8276, 2012.

Fischer, M., Huss, M., Barboux, C., and Hoelzle, M.: The New Swiss Glacier Inventory SGI2010: Relevance of Using High-Resolution Source Data in Areas Dominated by Very Small Glaciers, Arct. Antarct. Alp. Res., 46, 933–945, https://doi.org/10.1657/1938-4246-46.4.933, 2014.

Freudiger, D., Mennekes, D., Seibert, J., and Weiler, M.: Historical glacier outlines from digitized topographic maps of the Swiss Alps, Earth Syst. Sci. Data, 10, 805–814, https://doi.org/10.5194/essd-10-805-2018, 2018.

Frick, C., Steiner, H., Mazurkiewicz, A., Riediger, U., Rauthe, M., Reich, T., and Gratzki, A.: Central European high-resolution gridded daily data sets (HYRAS): Mean temperature and relative humidity, Meteorol. Z., 23, 15–32, https://doi.org/10.1127/0941-2948/2014/0560, 2014.

Gennaretti, F., Sangelantoni, L., and Grenier, P.: Toward daily climate scenarios for Canadian Arctic coastal zones with more realistic temperature-precipitation interdependence, J. Geophys. Res.-Atmos., 120, 11862–11877, https://doi.org/10.1002/2015JD023890, 2015.

Gobiet, A., Kotlarski, S., Beniston, M., Heinrich, G., Rajczak, J., and Stoffel, M.: 21st century climate change in the European Alps – A review, Sci. Total Environ., 493, 1138–1151, https://doi.org/10.1016/j.scitotenv.2013.07.050, 2014.

Godsey, S. E., Kirchner, J. W., and Tague, C. L.: Effects of changes in winter snowpacks on summer low flows. Case studies in the Sierra Nevada, California, USA, Hydrol. Process., 28, 5048–5064, https://doi.org/10.1002/hyp.9943, 2014.

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., and Engen-Skaugen, T.: Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – a comparison of methods, Hydrol. Earth Syst. Sci., 16, 3383–3390, https://doi.org/10.5194/hess-16-3383-2012, 2012.

Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Hagg, W., Braun, L. N., Kuhn, M., and Nesgaard, T. I.: Modelling of hydrological response to climate change in glacierized Central Asian catchments, J. Hydrol., 332, 40–53, https://doi.org/10.1016/j.jhydrol.2006.06.021, 2007.

Hakala, K., Addor, N., and Seibert, J.: Hydrological modeling to evaluate climate model simulations and their bias correction J. Hydrometeorol., 19, 1321–1337, https://doi.org/10.1175/JHM-D-17-0189.1, 2018.

Hänggi, P. and Weingartner, R.: Variations in Discharge Volumes for Hydropower Generation in Switzerland, Water Resour. Manage., 26, 1231–1252, https://doi.org/10.1007/s11269-011-9956-1, 2012.

Hanzer, F., Förster, K., Nemec, J., and Strasser, U.: Projected cryospheric and hydrological impacts of 21st century climate change in the Ötztal Alps (Austria) simulated using a physically based approach, Hydrol. Earth Syst. Sci., 22, 1593–1614, https://doi.org/10.5194/hess-22-1593-2018, 2018.

Hempel, S., Frieler, K., Warszawski, L., Schewe, J., and Piontek, F.: A trend-preserving bias correction – the ISI-MIP approach, Earth Syst. Dynam., 4, 219–236, https://doi.org/10.5194/esd-4-219-2013, 2013.

Hock, R.: Glacier melt: A review of processes and their modelling, Prog. Phys. Geogr., 29, 362–391, https://doi.org/10.1191/0309133305pp453ra, 2005.

Hottelet, C., Braun, L. N., Leibundgut, C., and Rieg, A.: Simulation of snowpack and discharge in an alpine karst basin, IAHS Publ., 218, 249–260, 1993.

Huss, M. and Farinotti, D.: Distributed ice thickness and volume of all glaciers around the globe, J. Geophys. Res., 117, F04010, https://doi.org/10.1029/2012JF002523, 2012.

Huss, M. and Hock, R.: Global-scale hydrological response to future glacier mass loss, Nat. Clim. Change, 8, 135–140, https://doi.org/10.1038/s41558-017-0049-x, 2018.

Huss, M., Jouvet, G., Farinotti, D., and Bauder, A.: Future high-mountain hydrology: A new parameterization of glacier retreat, Hydrol. Earth Syst. Sci., 14, 815–829, https://doi.org/10.5194/hess-14-815-2010, 2010.

Ivanov, M. A. and Kotlarski, S.: Assessing distribution-based climate model bias correction methods over an alpine domain:

Added value and limitations, Int. J. Climatol., 37, 2633–2653, https://doi.org/10.1002/joc.4870, 2017.

Jansson, P., Hock, R., and Schneider, T.: The concept of glacier storage: A review, J. Hydrol., 282, 116–129, https://doi.org/10.1016/S0022-1694(03)00258-0, 2003.

Jenicek, M., Seibert, J., and Staudinger, M.: Modeling of Future Changes in Seasonal Snowpack and Impacts on Summer Low Flows in Alpine Catchments, Water Resour. Res., 54, 538–556, https://doi.org/10.1002/2017WR021648, 2018.

Köplin, N., Viviroli, D., Schädler, B., and Weingartner, R.: How does climate change affect mesoscale catchments in Switzerland? – a framework for a comprehensive assessment, Adv. Geosci., 27, 111–119, https://doi.org/10.5194/adgeo-27-111-2010, 2010.

Köplin, N., Rößler, O., Schädler, B., and Weingartner, R.: Robust estimates of climate-induced hydrological change in a temperate mountainous region, Climatic Change, 122, 171–184, https://doi.org/10.1007/s10584-013-1015-x, 2014.

Kotlarski, S., Zubler, E., Fischer, A., Winter, K. J. P., Gobiet, A., and Liniger, M. A.: Patterns of 21st Century Climate Change in the European Alps: The CORDEX RCM ensembles, Geophysical Research Abstracts, Vol. 18, EGU2016-6247, 2016.

Li, C., Sinha, E., Horton, D. E., Diffenbaugh, N. S., and Michalak, A. M.: Joint bias correction of temperature and precipitation in climate model simulations, J. Geophys. Res.-Atmos., 119, 13153–13162, https://doi.org/10.1002/2014JD022514, 2014.

Lindström, G.: A simple automatic calibration routine for the HBV model, Hydrol. Res., 28, 153–168, 1997.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergström, S.: Development and test of the distributed HBV-96 hydrological model, J. Hydrol., 201, 272–288, https://doi.org/10.1016/S0022-1694(97)00041-3, 1997.

Maisch, M., Wipf, A., Denneler, B., Battaglia, J., and Benz, C.: Die Gletscher der Schweizer Alpen: Gletscherhochstand 1850. Aktuelle Vergletscherung, Gletscherschwundszenarien, vdf Hochschulverlag, Zurich, 2000.

Maraun, D.: Bias Correction, Quantile Mapping, and Downscaling: Revisiting the Inflation Issue, J. Climate, 26, 2137–2143, https://doi.org/10.1175/JCLI-D-12-00821.1, 2013.

Maraun, D.: Bias Correcting Climate Change Simulations – a Critical Review, Curr. Clim. Change Rep., 2, 211–220, https://doi.org/10.1007/s40641-016-0050-x, 2016.

Maraun, D. and Widmann, M.: Cross-validation of bias-corrected climate simulations is misleading, Hydrol. Earth Syst. Sci., 22, 4867–4873, https://doi.org/10.5194/hess-22-4867-2018, 2018.

Maraun, D., Shepherd, T. G., Widmann, M., Zappa, G., Walton, D., Gutiérrez, J. M., Hagemann, S., Richter, I., Soares, P. M. M., Hall, A., and Mearns, L. O.: Towards process-informed bias correction of climate change simulations, Nat. Clim. Change, 7, 664–773, https://doi.org/10.1038/nclimate3418, 2017.

Mehrotra, R. and Sharma, A.: Correcting for systematic biases in multiple raw GCM variables across a range of timescales, J. Hydrol., 520, 214–223, https://doi.org/10.1016/j.jhydrol.2014.11.037, 2015.

Mehrotra, R. and Sharma, A.: A Multivariate Quantile-Matching Bias Correction Approach with Auto- and Cross-Dependence across Multiple Time Scales: Implications for Downscaling, J. Climate, 29, 3519–3539, https://doi.org/10.1175/JCLI-D-15-0356.1, 2016.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J.-F., Matsumoto, K., Montzka, S. A., Raper, S. C. B., Riahi, K., Thomson, A., Velders, G. J. M., and van Vuuren, D. P. P.: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, Climatic Change, 109, 213–241, https://doi.org/10.1007/s10584-011-0156-z, 2011.

Messerli, B., Viviroli, D., and Weingartner, R.: Mountains of the World: Vulnerable Water Towers for the 21st Century, Ambio, 2004, 29–34, 2004.

Milano, M., Reynard, E., Bosshard, N., and Weingartner, R.: Simulating future trends in hydrological regimes in Western Switzerland, J. Hydrol.: Reg. Stud., 4, 748–761, https://doi.org/10.1016/j.ejrh.2015.10.010, 2015.

Moore, R. D., Fleming, S. W., Menounos, B., Wheate, R., Fountain, A., Stahl, K., Holm, K., and Jakob, M.: Glacier change in western North America: Influences on hydrology, geomorphic hazards and water quality, Hydrol. Process., 23, 42–61, https://doi.org/10.1002/hyp.7162, 2009.

Müller, F., Caflish, T., and Müller, G.: Firn und Eis der Schweizer Alpen: Gletscherinventar, vdf Hochschulverlag, Zurich, 1976.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I – A discussion of principles, J. Hydrol., 10, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

OSHD-SLF: Datenblatt zur SWE-Kartenserie 1972–2012 // KHR-Projekt, Stand 30.11.2013, Operationeller Schneehydrologischer Dienst des SLF (OSHD-SLF), Davos, 2013.

Oudin, L., Hervieu, F., Michel, C., Perrin, C., Andréassian, V., Anctil, F., and Loumagne, C.: Which potential evapotranspiration input for a lumped rainfall–runoff model? Part 2 – Towards a simple and efficient potential evapotranspiration model for rainfall–runoff modelling, J. Hydrol., 303, 290–306, https://doi.org/10.1016/j.jhydrol.2004.08.026, 2005.

Paul, F., Frey, H., and Le Bris, R.: A new glacier inventory for the European Alps from Landsat TM scenes of 2003: challenges and results, Ann. Glaciol., 52, 144–152, https://doi.org/10.3189/172756411799096295, 2011.

Pellicciotti, F., Bauder, A., and Parola, M.: Effect of glaciers on streamflow trends in the Swiss Alps, Water Resour. Res., 46, W10522, https://doi.org/10.1029/2009WR009039, 2010.

Pitié, F., Kokaram, A. C., and Dahyot, R.: Automated colour grading using colour distribution transfer, Comput. Vis. Image Understand., 107, 123–137, https://doi.org/10.1016/j.cviu.2006.11.011, 2007.

Räty, O., Räisänen, J., Bosshard, T., and Donnelly, C.: Intercomparison of Univariate and Joint Bias Correction Methods in Changing Climate From a Hydrological Perspective, Climate, 6, 33, https://doi.org/10.3390/cli6020033, 2018.

Rauthe, M., Steiner, H., Riediger, U., Mazurkiewicz, A., and Gratzki, A.: A Central European precipitation climatology – Part I: Generation and validation of a high-resolution gridded daily data set (HYRAS), Meteorol. Z., 22, 235–256, https://doi.org/10.1127/0941-2948/2013/0436, 2013.

R Core Team: R: A language and environment for statistical computing, R Foundation for Statistical Computing, Vienna, Austria, available at: https://www.R-project.org/ (last access: 4 March 2019), 2018.

Schaefli, B., Manso, P., Fischer, M., Huss, M., and Farinotti, D.: The role of glacier retreat for Swiss hy-

dropower production, Renewable Energy, 132, 615–627, https://doi.org/10.1016/j.renene.2018.07.104, 2019.

Seibert, J.: Multi-criteria calibration of a conceptual runoff model using a genetic algorithm, Hydrol. Earth Syst. Sci., 4, 215–224, https://doi.org/10.5194/hess-4-215-2000, 2000.

Seibert, J. and Vis, M. J. P.: Teaching hydrological modeling with a user-friendly catchment-runoff-model software package, Hydrol. Earth Syst. Sci., 16, 3315–3325, https://doi.org/10.5194/hess-16-3315-2012, 2012.

Seibert, J., Vis, M. J. P., Kohn, I., Weiler, M., and Stahl, K.: Technical note: Representing glacier geometry changes in a semi-distributed hydrological model, Hydrol. Earth Syst. Sci., 22, 2211–2224, https://doi.org/10.5194/hess-22-2211-2018, 2018.

Sevruk, B.: Reliability of precipitation measurement, in: Precipitation measurements: WMO/IAHS/ETH Workshop on Precipitation Measurements, edited by: Sevruk, B., 3–7 December 1989, St. Moritz, Zurich, 13–19, 1989.

Stahl, K., Weiler, M., Freudiger, D., Kohn, I., Seibert, J., Vis, M., Gerlinger, K., and Böhm, M.: The snow and glacier melt components of streamflow of the river Rhine and its tributaries considering the influence of climate change, Final report to the International Commission for the Hydrology of the Rhine Basin (CHR), available at: https://www.chr-khr.org/en/publications (last access: 5 June 2018), 2017.

Teutschbein, C. and Seibert, J.: Bias correction of regional climate model simulations for hydrological climate-change impact studies: Review and evaluation of different methods, J. Hydrol., 456–457, 12–29, https://doi.org/10.1016/j.jhydrol.2012.05.052, 2012.

Vrac, M.: Multivariate bias adjustment of high-dimensional climate simulations: The Rank Resampling for Distributions and Dependences ($R^2D^2$) bias correction, Hydrol. Earth Syst. Sci., 22, 3175–3196, https://doi.org/10.5194/hess-22-3175-2018, 2018.

Weiler, M., Seibert, J., and Stahl, K.: Magic components – why quantifying rain, snowmelt, and icemelt in river discharge is not easy, Hydrol. Process., 32, 160–166, https://doi.org/10.1002/hyp.11361, 2018.

Wilcke, R. A. I., Mendlik, T., and Gobiet, A.: Multi-variable error correction of regional climate models, Climatic Change, 120, 871–887, https://doi.org/10.1007/s10584-013-0845-x, 2013.

Zscheischler, J., Fischer, E. M., and Lange, S.: The effect of univariate bias adjustment on multivariate hazard estimates, Earth Syst. Dynam., 10, 31–43, https://doi.org/10.5194/esd-10-31-2019, 2019.

PAPER IV

Regional Environmental Change (2020) 20:34
https://doi.org/10.1007/s10113-020-01615-8

ORIGINAL ARTICLE

# Sensitivity of discharge projections to potential evapotranspiration estimation in Northern Tunisia

Hamouda Dakhlaoui[1,2,3] · Jan Seibert[3,4] · Kirsti Hakala[3,5]

Received: 7 August 2018 / Accepted: 23 November 2019


## Abstract

Tunisia has a long history of coping with water scarcity, and the quantification of climate change impacts on runoff is important for future water management. A major requirement for such studies is an estimation of potential evapotranspiration (PET), which is challenging as many regions often lack the observational data needed for physically based PET equations. In this study, different PET estimation approaches were used to study the impact of PET estimation on discharge projections for catchments in Northern Tunisia. Discharge was simulated for five catchments using three rainfall-runoff models (RRMs): HBV, GR4 and IHACRES. A general differential split sample test (GDSST) was used for an RRM robustness evaluation based on subperiods with contrasting climatic conditions for the 1970–2000 period. Three cases with varying PET were considered: (1) daily calculated PET, (2) long-term daily mean PET with the same values for calibration and validation periods (calculated over the calibration period) and (3) long-term daily mean PET varying between calibration and validation periods (calculated over the calibration and validation period separately). Over the historical period, the comparison between cases 1 and 3 showed little impact of reduced PET information on the RRM performance and robustness. The comparison of cases 2 and 3 indicated a limited impact of varying PET between calibration and validation on the RRM results. The impact of varying levels of PET information on hydrological projections was also analysed over two future 30-year periods: mid-term period (2040–2070) and long-term period (2070–2100), with two representative concentration pathway scenarios (RCPs 4.5 and 8.5), by comparing cases 1 and 2. The projected discharge with constant PET (case 2) was generally lower than the projected discharge with variable PET (case 1) but the difference in volume change did not exceed 9% for both the time period and the RCP scenario considered. While PET slightly increased under the different climate change scenarios, actual evapotranspiration (AET) was found to decrease. These opposite trends of PET and AET can be attributed to the projected decrease in precipitation. Overall, our results demonstrate that discharge, in semi-arid regions like Northern Tunisia, is not sensitive to PET estimates since AET is mainly controlled by the availability of soil moisture. This finding is useful for performing studies of climate change impact on hydrological cycles in arid regions, as our study shows that simple PET estimation is a valid approach for such studies.

**Keywords** Rainfall-runoff modelling · PET · GDSST · Hydrological projection · EURO-CORDEX · Tunisia

This article is part of the Topical Collection on *Climate change impacts in the Mediterranean*

**Electronic supplementary material** The online version of this article (https://doi.org/10.1007/s10113-020-01615-8) contains supplementary material, which is available to authorized users.

✉ Hamouda Dakhlaoui
hammouda.dakhlaoui@laposte.net

✉ Jan Seibert
jan.seibert@slu.se; jan.seibert@geo.uzh.ch

Kirsti Hakala
kirsti.hakala@unimelb.edu.au

[1] LMHE, Ecole Nationale d'Ingénieurs de Tunis, Université Tunis El Manar, BP 37 Le Belvedère, 1002 Tunis, Tunisia

[2] Ecole Nationale d'Architecture et d'Urbanisme, Université de Carthage, Sidi Bou Said, Tunisia

[3] Department of Geography, University of Zurich, Zürich, Switzerland

[4] Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, Uppsala, Sweden

[5] Department of Infrastructure Engineering, University of Melbourne, Victoria, Australia

Published online: 05 March 2020

## Introduction

Tunisia has a long history of coping with water scarcity, characterized by extended periods of drought as well as persistent issues with water quality and over-extraction. Northern Tunisia produces approximately 83% of the surface water in Tunisia while providing several water transfers to the eastern and the southern parts of the country (ITES 2014). Recent climate change scenarios project around a 20% decrease in total precipitation and a + 1 °C to + 3 °C increase in mean annual temperature by 2050 compared with the 1971–1990 period (Terink et al. 2013; Tramblay et al. 2017). This could result in a substantial decrease in fresh water availability in the future, which would have a dramatic impact on various socio-economic sectors (e.g. agriculture and tourism), and implies significant risks for ecosystems and for human well-being (Cramer et al. 2018). Adaptation to climate change is therefore crucial, and this level of susceptibility calls for an understanding of local climate change impacts on water resources. Rainfall-runoff models (RRMs), forced by regional climate models, are widely used to assess the hydrological impacts of climate change at the catchment scale (e.g. Ruelland et al. 2015; Hakala et al. 2019). Evapotranspiration is an essential part of the water balance at the catchment scale, especially for the Mediterranean region where around 90% of the annual rainfall can be lost through evapotranspiration (Wilcox et al. 2003). Therefore, to accurately simulate discharge, the estimation of potential evapotranspiration (PET) is required as an input. However, studies that evaluated the hydrological impacts of climate change on water resources in the Southwest Mediterranean Rim region (e.g. Ruelland et al. 2015; Sellami et al. 2015; Marchane et al. 2017) did not consider the sensitivity of PET estimation on hydrological projections.

Although a bias in PET estimation may be compensated during the calibration process by model parameters (Oudin et al. 2005a), it can have a considerable impact for periods outside calibration, especially in the case of climate variability. This, in turn, can lead to significant errors when modelling hydrological peaks and recession characteristics (Andréassian et al. 2004). The sensitivity of discharge to PET estimation may depend on the climate conditions, with a corresponding greater sensitivity within wet regions (Sperna Weiland et al. 2012; Seiller and Anctil 2014) than dry regions (Sheffield et al. 2012; Kingston et al. 2009). Recently, Guo et al. (2017) found that discharge sensitivity to PET in five catchments in climatologically different regions of Australia depends on both PET estimation and the methods used by hydrological models to convert PET to actual evapotranspiration (AET).

PET values must be expected to vary for a changing climate (Prudhomme and Williamson 2013; Seiller and Anctil 2014), and the estimation of PET could be an additional source of uncertainty for hydrological projections. Several studies show that the impacts of climate change on discharge depend on the PET formula used (Seiller and Anctil 2016; Bae et al. 2011; Sperna Weiland et al. 2012). Physically-based PET formulas could theoretically have the potential to provide more accurate PET estimations compared with temperature-based formulas because they consider the different climate drivers that affect the evaporative demand (Prudhomme and Williamson 2013). Assuming that anthopogenic $CO_2$ emissions will keep on influencing temperatures over the next century (IPCC 2013), the use of temperature-based formulas results in a substantial increase in projected PET. However, several studies have investigated the pan evaporation paradox (Li et al. 2013) and have shown that future potential evaporation could be attenuated by the change in other climate variables such as the expected increase of air humidity (Wang et al. 2017) and decreasing trend for wind speed (Mansour et al. 2017).

Although there are inherent limitations to using temperature-based PET formula, it is often not feasible to alternatively use a physically based formula due to a large number of climate variables required to fulfil such equations and a lack of observational data. In this regard, data availability is a particular limitation in regions such as Northern Tunisia due to limited climate records (e.g. Jabloun and Sahli 2008; Aouissi et al. 2016). The use of temperature-based PET formulas is, thus, often considered the only possible solution. It is therefore not surprising that the hydrological projection studies performed over Northwest Africa (Maghreb region) have mainly utilized temperature-based formulas (Marchane et al. 2017; Tramblay et al. 2016; Ruelland et al. 2015). Given the widespread use of such simple formulas, the impact of the use of temperature-based PET formulas on hydrological projections, rather than physically based PET formulas, needs to be explored in this region of the world. For this purpose, it would, in theory, be ideal to compare the hydrological projections generated with hydrological models forced by temperature-based PET formula to those forced by physically based PET formulas. However, in the context of data scarcity, such a comparison is not feasible. Therefore, this study employs a novel methodology, with low data requirements, to test the sensitivity of hydrological projections to PET estimation.

The main goal of the present study is to provide a methodology for the assessment of the potential impact of PET estimation on the prediction capacity of RRMs and on hydrological projections, within a data-scarce region. Additionally, this study aims to provide a first quantification on the impact of climate change on discharge in Northern Tunisia, and thereby encourage additional climate change impact studies for Tunisia and regions with similar climate and data availability. To this end, this study addresses the three following research questions:

- What is the sensitivity of RRM to PET estimation in an arid region?
- What are the likely impacts of climate change on the hydroclimatology of Northern Tunisia?
- How sensitive are the trajectories of hydrological projections to PET estimation?

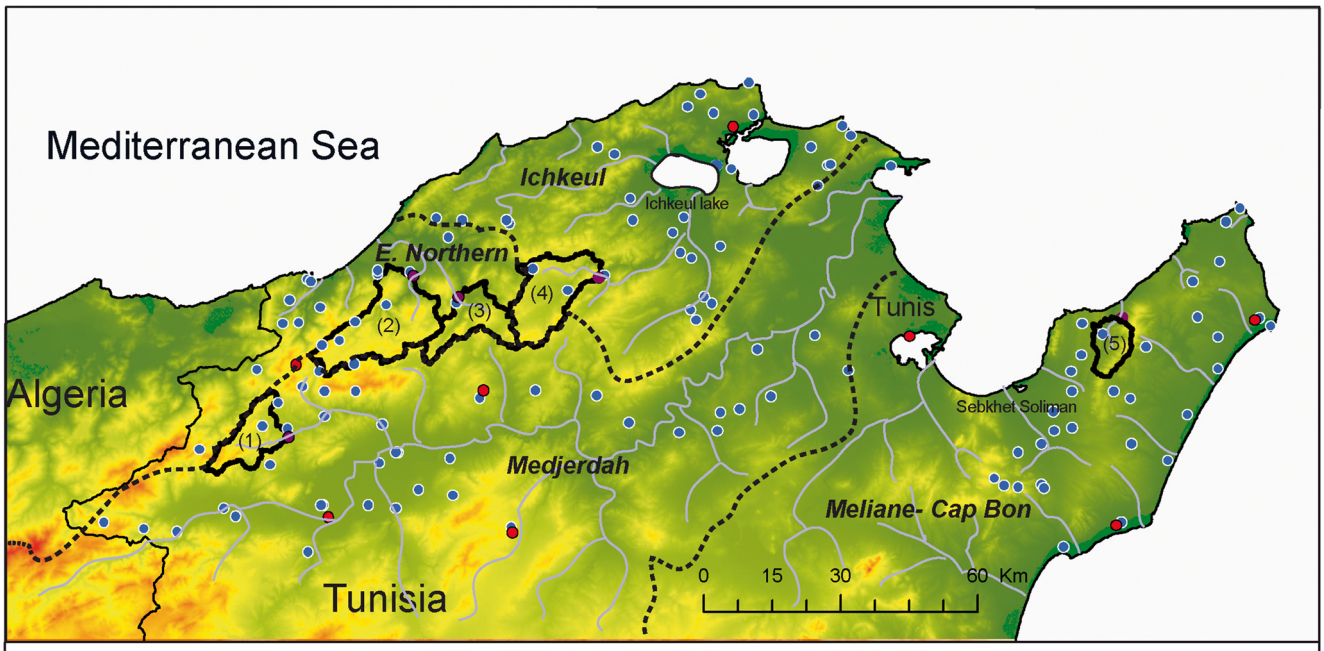
## Data and models

### Study areas

Five catchments located in Northern Tunisia were selected for this study (Fig. 1). The discharge can be considered 'natural' since these catchments are located upstream from major hydraulic installations, such as dams and water transfers. The study catchments are situated within a semi-arid to humid Mediterranean climate with a warm and dry season which extends from June to August (Henia 2008). When averaged over the catchment areas, PET always exceeds precipitation (Fig. 1). In this region, actual evapotranspiration is mainly limited by water availability for most parts of the year (Tramblay et al. 2017). The catchments are located in a pivotal area, which serves a strategic role as a water supplier for the rest of the country (Ben Fraj et al. 2019). Some of the streams feed into wetlands (Lake Ichkeul, Sebkhet Soliman, etc). Lake Ichkeul, and its marshes, is one of the most important wetlands of the western Mediterranean basin and is an essential stopover point for birds migrating between Eurasia and Sub-Saharan Africa (Hamdi et al. 2012).

### Hydrological models

Discharge simulations have been shown to be sensitive to the methods used by hydrological models to convert PET to AET (Guo et al. 2017). This encouraged us to use different model structures for this study. Three simple bucket-type RRMs, all running at the daily time step, were used: GR4J (Perrin et al. 2003), HBV (Lindström et al. 1997) and IHACRES (Jakeman et al. 1990). These models differ in the way they conceptualise hydrological processes and in their complexity: GR4J has four parameters, whereas HBV has eight and IHACRES six. These models differ in how AET is computed from PET, which allows for an interesting inter-comparison of model simulations. In GR4, a production function accounts for precipitation and potential evapotranspiration and determines the effective precipitation that contributes to flow and supplies the




| N° | Catchment | Outlet | DGRE* Code | Area (Km$^2$) | Av. Elev. (m) | Precip. (mm) | Temp. (°C) | PET (mm) | Runoff (mm) |
|---|---|---|---|---|---|---|---|---|---|
| (1) | Rhezala | Fernana | 5303510 | 138 | 470 | 857 | 16.3 | 1020 | 101 |
| (2) | Melah | Ouchtata | 3700250 | 315 | 423 | 928 | 16.5 | 1023 | 179 |
| (3) | Maaden | Bou Brima | 3700350 | 145 | 337 | 783 | 16.6 | 1029 | 135 |
| (4) | Joumine | J. Antra | 3600110 | 234 | 375 | 742 | 16.3 | 1016 | 116 |
| (5) | El Abid | Pont-route | 4300150 | 81 | 202 | 530 | 17.6 | 1061 | 40 |

**Fig. 1** Locations of study catchments are indicated by a bold black line along with the locations of the precipitation (blue circles), temperature (red circles) and stream flow gauges (purple circles). The main hydro-climatic characteristics are averaged over the period 1970–2000

reservoir. AET is calculated according to a parabolic function of soil moisture. Only one free parameter is dedicated to the production module. In HBV, AET equals PET when the soil storage is filled to a certain fraction (as specified by a model parameter) and decreases linearly for smaller soil water storage values. In IHACRES, the input rainfall is partitioned explicitly into drainage, evapotranspiration, and changes in catchment moisture. AET is calculated according to an exponential function of soil moisture. IHACRES allocates two free parameters to the production module. All three models have recently been used successfully to simulate discharge from other catchments in Tunisia (Bargaoui et al. 2008; Dakhlaoui et al. 2009, 2012; Abbaris et al. 2014). Additional information for each model can be found in Tab. S1 (Online Resources). The model parameters were calibrated using the Shuffle Complex Evolution algorithm (Duan et al. 1992) and the Kling-Gupta Efficiency (KGE, Gupta et al. 2009) as objective functions. The hydrological models were run at a daily time step, but their efficiency evaluation was performed on 10-day averages, since our study primarily focuses on water resources rather than on the day-to-day variation and the averaged discharge values were more reliable than the daily values. Model performance during validation was evaluated based on the Nash-Sutcliffe Efficiency (NSE, Nash and Sutcliffe 1970) and the relative volume error (VE) in mean discharge.

### In situ hydro-climatic data

In total, 123 rain gauges situated in the study area were used. The selected gauges have some missing data but do not exceed 30% of the total data over the period of 1970–2000, thus providing a stable, coherent network of measurements for the spatial interpolation of rainfall forcing (see Fig. 1). Eight meteorological stations with daily temperature data were used. Climate forcing data was generated by spatially interpolating the station data, on a 2-km horizontal grid spacing, using an inverse distance weighting technique. Based on the available data, the approach of Oudin et al. (2005b) was chosen to estimate PET. This formula is based on estimated clear daily sky solar radiation and mean daily air temperature. It is expressed as follows:

$$E_{\text{pot}} = \frac{R_e}{\lambda \rho}\frac{T_a + 5}{100} \quad \text{if } T_a > -5$$
$$E_{\text{pot}} = 0 \quad \text{otherwise} \tag{1}$$

where $E_{\text{pot}}$ is the rate of potential evapotranspiration (mm day $^{-1}$); $R_e$ is extraterrestrial radiation (MJ m$^{-2}$ day$^{-1}$); $\lambda$ is the latent heat flux (MJ kg$^{-1}$); $\rho$ is the density of water (kg m$^{-3}$) and $T_a$ is mean air temperature (°C).

### High-resolution climate simulations and bias correction of climate variables

Different daily temperature and precipitation time series, simulated by eleven GCM-RCMs (general circulation models forcing regional climate models), were obtained from the Coordinated Regional Downscaling Experiment (CORDEX; www.cordex.org). Given the location of the catchments, GCM-RCMs from the European domain were selected (EURO-CORDEX, http://www.euro-cordex.net/). EURO-CORDEX simulations are the most recent high-resolution climate projections for the European domain with a 0.11° resolution (~ 12.5-km horizontal grid spacing). The GCM-RCM historical simulations span over the period 1970–2005, and the future period was divided over two periods: mid-term (2040–2070) and long-term (2070–2100). Two representative concentration pathway (RCP) scenarios were considered for the projections: RCP 4.5 and RCP 8.5 (IPCC, 2013). RCP 4.5 assumes that global annual greenhouse gas emissions peak between around 2040, then decline thereafter. In RCP 8.5, emissions continue to rise throughout the twenty-first century (Meinshausen et al. 2011).

## Methods

### Impact of potential evapotranspiration estimation on rainfall-runoff model performance and robustness

To evaluate the sensitivity of hydrological projections to PET estimation, we first assessed the sensitivity of RRM to PET estimation under historical climate conditions, using a differential split sample test (DSST; Klemeš 1986). This step is needed so that we may check whether it is reasonable to extrapolate our methods to future period. We propose an evaluation of the sensitivity of discharge simulations to PET formulation, based on a comparison of three cases of PET estimations (see Fig. S1 in Online Resources):

- Case 1: Daily PET calculated over the calibration and validation period separately (without averaging).
- Case 2: Long-term daily mean PET with the same values for calibration and validation periods (PET calculated over the calibration period and then averaged to create a series of 366 values, which were repeated for both the calibration and validation period).
- Case 3: Long-term daily mean PET varying between calibration and validation periods (PET calculated over the calibration period and validation period separately. The two sets of daily values were then averaged to create two sets of 366 values, which were then used for calibration and validation separately).

Over the historical period, the comparison of RRM performance and robustness between cases 1 and 3 allowed for an evaluation of the impact of long-term averaging of PET on the hydrological model performance and robustness. The comparison of these two cases was also chosen in order to indicate whether the long-term averaging of PET was acceptable for hydrological projections. The comparison of cases 2 and 3 allowed for the evaluation of the impact of keeping PET the same versus allowing it to change between calibration and validation periods on RRM performance.

The evaluation of RRM robustness was performed using DSST (Klemeš 1986). DSST is the typical method to investigate RRM robustness under climate variability (Seibert 2003; Hartmann and Bárdossy 2005; Fowler et al. 2016; Dakhlaoui et al. 2017; Vormoor et al. 2018). It consists of calibration and validation exercises of RRM under climate-contrasted conditions. The idea behind performing a DSST is that the errors made by extrapolation from certain observed climate conditions to different observed conditions might correspond to the errors made when using observed data for calibration and extrapolating this to future climatic conditions (Seibert 2003). Dakhlaoui et al. (2019) proposed a generalization of DSST called general differential split sample test (GDSST) based on an oriented bootstrap applied to discontinuous subperiods, which allows for a large number of calibration and validation exercises and a better sample spread, with more contrasted subperiods (in precipitation and temperature) compared with previous techniques. This technique helps with the evaluation of RRM robustness for a broad range of climate conditions. We used the GDSST as proposed by Dakhlaoui et al. (2019) to generate 100 climate-contrasted subperiods (precipitation and temperature) over the observed period 1970–2000 for each catchment. The duration of each subperiod was 15 years, which could be non-consecutive. We used each subperiod as a calibration period for the RRM and evaluated the performance of hydrological predictions in the complementary subperiod. In contrast to other studies, which used continuous calibration and validation periods for the DSST (e.g. Coron et al. 2012; Brigode et al. 2013), we used discontinuous subperiods but always used entire hydrological years. The initial conditions for each hydrological year were checked so that they are not significantly affected by the climatic conditions of the preceding year. It should be noted that the climate of Northern Tunisia has a long dry season (June to August) during which most of the soil moisture is evaporated resulting in almost the same (dry) initial conditions in the following hydrological year.

We performed the calibration and validation exercise for each catchment for the 100 GDSST subperiods for cases 1, 2 and 3 separately. The model transferability was evaluated by the change in NSE and VE between the calibration and validation periods. NSE was based on a ratio between the squared model error and the variance of observed flows. Hence, any changes in variance or volumes between contrasted climatic periods such as dry/wet could have an impact on the comparison of results. Performance measures like NSE for different simulation time periods cannot easily be compared. Therefore, we evaluated model transferability by calculating the differences between NSE resulting from calibration period (receiver) and the NSE calculated over the same period but with parameters provided by model calibration on validation subperiods (donor). The same was done for VE. We evaluated model transferability in terms of NSE and VE as a function of the change in total precipitation and mean temperature ($\Delta P$ and $\Delta T$) between the validation and calibration periods. To study the origin of the differences in RRM performances between the different cases, we also looked at the differences in terms of PET and AET between the calibration and validation periods.

## Projections of climate change impacts

Quantile mapping (QM) was utilized as a bias correction method to correct the daily precipitation and temperature of the GCM-RCMs. Previous studies have shown QM to outperform other bias correction methods (Teutschbein and Seibert 2012; Chen et al. 2013). The aim of QM is to correct the distribution of the climate model data so that it matches the distribution of the observational data. It consists of estimating quantiles for both observation and modelled climate variable under a control period. A transfer function is then created by interpolation between corresponding quantile values, which is applied to the projected climate variable. Here, the cumulative distribution functions (CDFs) of observed and modelled climate variable were estimated using empirical percentiles. Values in between the percentiles were approximated using linear interpolation. In cases where new GCM-RCM values (such as from the projected period) were larger than the control values used to estimate the empirical CDF, a linear regression fit was used to extrapolate beyond the range of observations. The transfer function of the QM was based on a 30-year control period (1970–2000), which is the recommended time length for climate applications (WMO, 2011). The transfer function of the QM was then applied to bias correct the GCM-RCMs projected daily precipitation and temperature over the mid-term (2040–2070) and long-term (2070–2100) future periods for the two RCPs. The RRMs were forced by the bias-corrected GCM-RCM climate data (precipitation and temperature) to generate the projected discharge.

## Evaluating the sensitivity of hydrological projections to potential evapotranspiration estimation

Two cases were used to compare the impact of PET estimation on hydrological projections: (i) case 1: PET calculated at the daily scale, using the Oudin formula, over the historical and future period. In case 1, the observed temperature is used for the historical period, and GCM-RCM bias-corrected temperature is used for the calculation of PET for the future periods. (ii) Case 2: The

long-term daily mean PET is calculated and used for RRM calibration over the historical period and also repeated over the future time periods.

Case 1 is meant to represent a situation where a simple temperature-based PET formula, which is only sensitive to change in temperature, is used. We hypothesised that this case would result in the highest future increase in PET, due to the increasing of temperatures found within the RCP scenarios. Case 2 represents a different situation, similar to the pan evaporation paradox, which expects an attenuation of future PET by the future change in other climatic drivers, until stabilisation. This kind of attenuation of future PET by other climatic drivers could be addressed only by a physically based PET formula. The comparison of these two extreme cases allows for the evaluation of the sensitivity of hydrological projections to the PET estimation and assessing the potential loss in hydrological projection caused by using a simple temperature-based PET formula rather than a more physically based PET formula.

We considered the variance of the hydrological projections as an estimate of their uncertainty and used an analysis of variance (ANOVA, Hawkins and Sutton 2009) technique to quantify the contribution of uncertainty stemming from different elements of the modelling chain to the total uncertainty. More specifically, this provides a sensitivity test of whether PET estimation contributes to the total uncertainty compared with the other sources of uncertainty. A similar analysis (using ANOVA) was performed by Addor et al. (2014) to quantify the contribution from different aspects of the modelling chain to hydrological projection uncertainty for six Swiss catchments; however, that study did not investigate PET estimation uncertainty. Following this example, and extending it to additionally take PET estimation uncertainty into account, six sources of uncertainty were compared: (1) emission scenarios (RCP 4.5 and RCP 8.5), (2) GCM-RCMs, (3) hydrological models (HBV, GR4 and IHACRES), (4) PET estimation (cases 1 and 2), (5) the sum of the significant interaction between factors and (6) residual error. The fraction of explained variance of the discharge change for the two 30-year future periods (2040–2070 and 2070–2100) compared with the reference period (1970–2000), by each factor, was computed by dividing the variance of each factor by the total variance. The analysis was performed for each catchment and each of the two future periods.

## Results

### Evaluating the effect of potential evapotranspiration estimation on rainfall-runoff model robustness

#### Effect of long-term potential evapotranspiration averaging on rainfall-runoff model robustness

Our results show that when daily PET (case 1) is compared with long-term daily mean PET (case 3), only minimal differences exist between the PET values. The biggest differences between case 1 and case 3 occur in summer, where discharge generation is at its lowest level during the year. Conversely, the smallest differences were observed in winter, where the generation of discharge is the greatest. A comparison between case 1 and case 3, over the period 1970–2000, for the O. Abid catchment is presented in the Online Resources (Fig. S2). Cases 1 and 3 exhibit similar performances over the calibration period in terms of NSE and VE (Online Resources, Fig. S3), which implies that temporally varying PET and a long-term daily mean PET produce similar discharge time series. Furthermore, this demonstrates the low sensitivity of hydrological models to PET interannual variation.

Figure 2 shows the mean results of GDSST experiments for each case of PET estimation, where $\Delta P$ and $\Delta T$ are compared for all catchments. Based on the NSE criterion, the transferability was gradually affected by the decrease in precipitation and the increase in temperature (Fig. 2a). Similarly, the transferability of VE was also more affected by changes in precipitation than in temperature (Fig. 2b). We observed no significant difference in RRM robustness between cases 1 and 3 (Fig. 2a, b). The results shown here are for the GR4 model. However, we observed a similar behaviour between the different models. (Results from HBV and IHACRES are shown in Fig. S4 within the Online Resources.)

To better understand the origin of the difference in RRM robustness between cases 1 and 3, Fig. 2 c and d show the difference in terms of PET and AET between calibration and validation for the GR4 model. The behaviour of PET is shown to be highly correlated to temperature variation, which is expected since temperature is the basis for the calculation of PET using the Oudin formula. However, this is not the case for AET, which is mainly dependent on water availability (i.e. precipitation). Instead, AET increases with a decrease in temperature (Fig. 3).

#### Effect of potential evapotranspiration estimation on the robustness of rainfall-runoff models

Similar to the results presented in the previous section, an additional experiment was performed to evaluate the change in RRM robustness between cases 2 and 3 (see Fig. 2a, b for GR4 and Fig. S3 (Online Resources) for HBV and IHACRES). Small differences were found between the robustness of the RRMs when comparing the two cases, in terms of VE (difference between − 4 and + 4%). For NSE, the difference between the two cases was around zero. The RRMs forced by long-term daily mean PET with the same values for calibration and validation periods tend to overestimate VE for positive $\Delta T$, and underestimate VE for negative $\Delta T$, compared with the case where PET varies between calibration and validation periods. Figure 2 c and d show the difference in terms of PET and AET between calibration and validation periods between cases

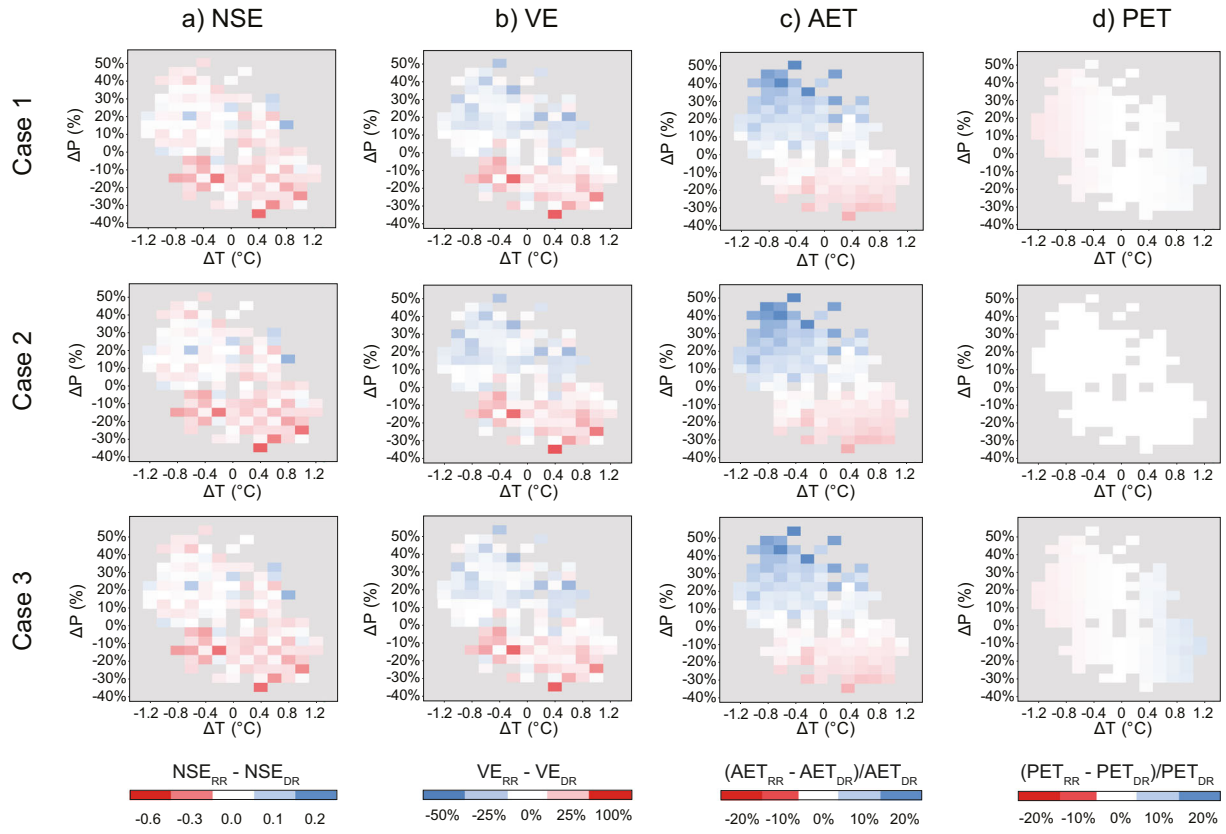



**Fig. 2** Evaluation of GR4 model efficiency as a function of change in precipitation and temperature (ΔP and ΔT). Each coloured square represents the mean results for the five catchments. Model transferability was evaluated according to differences in **a** Nash-Sutcliffe Efficiency (NSE) and **b** relative error in mean discharge (VE) between the receiver (RR, i.e. validation) and the donor (DR, i.e. calibration) periods. The difference in terms of potential evapotranspiration (PET) and actual evapotranspiration (AET) between calibration and validation period is presented in a similar way in **c** and **d**

2 and 3. The AET of the different contrasted subperiods depends mainly on precipitation, e.g. larger AET values are observed for the wetter subperiod and vice versa. AET is less dependent on temperature (Fig. 3).

## Projections under climate change scenarios

The climate projections over the study catchments show that for precipitation, over the mid-term period, the RCP 4.5 scenario results in a decrease of about − 8% in total precipitation and − 15% for RCP 8.5. For the long-term time period, a decrease of about − 16% of total precipitation is expected for RCP 4.5 and − 26% for RCP 8.5. These changes mainly occur during the wet season (November to April). For temperature, over the mid-term period, an increase of + 0.44 to + 2.3 °C is projected with the scenario RCP 4.5 and + 1.2 to + 2.6 °C with scenario RCP 8.5. For temperature, over the long term, an increase of + 1.2 to + 2.3 °C is projected within the scenario RCP 4.5, and + 2.1 to + 4 °C with the scenario RCP 8.5. For PET calculated by the Oudin formula using the projected temperature (Fig. 4c), in over the mid-term period, an increase of + 1.8 to + 9.6% is projected with the scenario RCP 4.5 and + 4.8 to + 11% with scenario RCP 8.5. For PET, over the long term, an increase of + 3.5 to + 11% is projected within the scenario RCP 4.5, and + 8.8 to + 17% with the scenario RCP 8.5. Contrary to precipitation, these changes are projected mainly during the summer months. PET changes are likely to have little impact on discharge since there is already very little discharge during the summer. In contrast, the decrease in precipitation amounts during winter months may have a critical impact on water resources. Our results also show that the climate change signal is very different from one GCM-RCM to another; however, all GCM-RCMs depict a warmer future climate and almost all climate models expect dryer conditions for the future (Fig. 4a, b). (Detailed results of the changes can be found within the Online Resources (Fig. S5).)

## Evaluating the sensitivity of hydrological projections to potential evapotranspiration estimation

The median of the hydrological projections for the Maaden catchment, when comparing cases 1 and 2 with three RRMs (Fig. 4d, detailed results for the remaining catchments are presented in Fig. S5 of the Online Resources), shows that, for the mid-term period, the RCP 4.5 scenario results in a

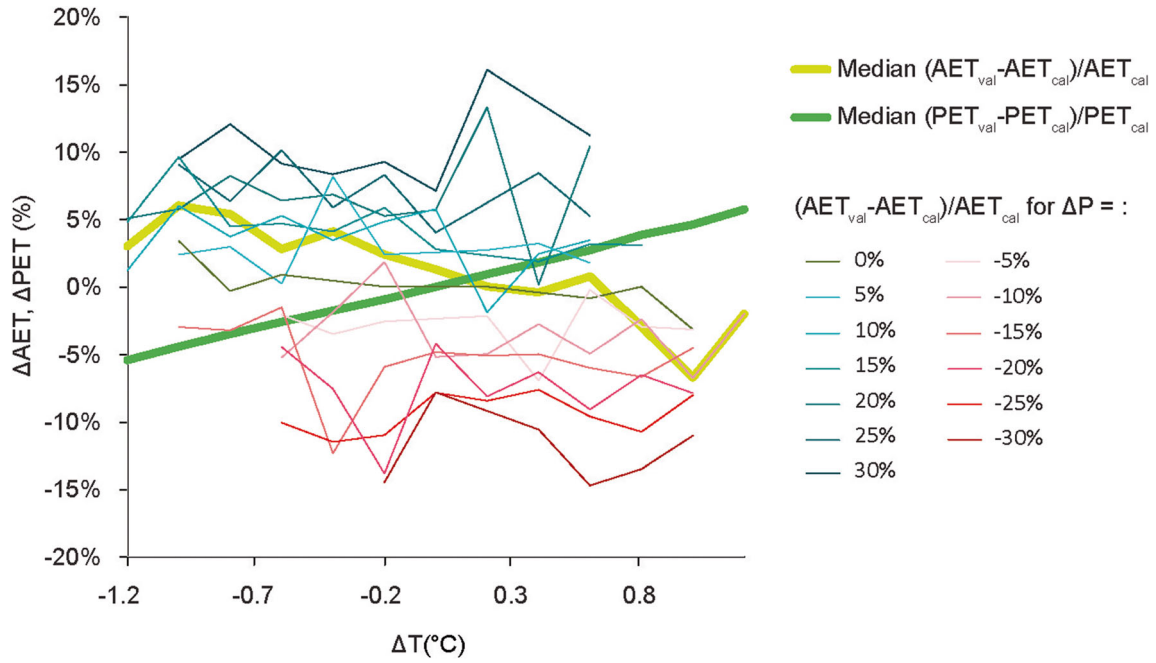



**Fig. 3** Change in actual evapotranspiration (AET) and potential evapotranspiration (PET) as a function of change in precipitation and temperature (ΔP and ΔT) between calibration (cal) and validation period (val) for the hydrological model GR4

decrease of discharge of 10 to 30% and 25 to 38% for the RCP 8.5 scenario. For the long-term period, a decrease of 20 to 37% of discharge is expected for RCP 4.5 and 41 to 58% for RCP 8.5.

However, for case 2, the estimated hydrological impact is by approximately − 2 to − 8% smaller than for case 1. The difference in mean annual discharge between the two cases is generally under 5% for the RCP 4.5 (for both the mid-term and long-tern periods) and also for RCP 8.5 for the mid-term period. The difference in volume change between cases 1 and 2 exceeds the 5% for RCP 8.5 over long-term periods, but it is always less than 10% (see Fig. S6 within the Online Resources). These results are robust, meaning that we find similar behaviour amongst the different RRMs.

The decomposition of the projection variance by ANOVA partitioning amongst the six sources of uncertainty shows that the GCM-RCMs are the dominant source of uncertainty. The difference between cases 1 and 2 represents only a small proportion of the variance in the hydrological projections (see Fig. S7 within Online Resources).

The hydrological response of Northern Tunisian catchments shows a high sensitivity to any change in precipitation. The evaluation of the elasticity of discharge to precipitation for the hydrological model GR4 (e.g. change in discharge

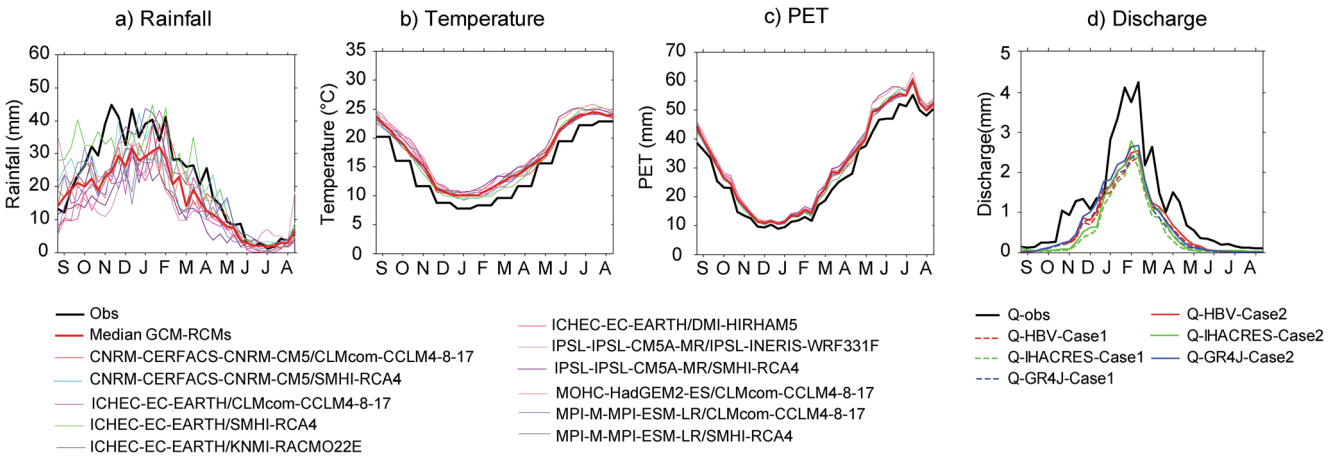



**Fig. 4** Projected (a) rainfall, (b) temperature, (c) potential evapotranspiration (PET) and (d) discharge (Q) for cases 1 and 2, from HBV, IHACRES and GR4J model, for the Maaden catchment, for representative concentration pathway scenario 8.5, over the long-term period (2070–2100) and compared with observation over the reference period 1970–2000 (Obs). GCM and RCM names are separated by a slash

between present and projected climate as a function of the change in precipitation) results in a hydrological multiplication factor of around two (Fig. 5, the results from the other models are included in the Online Resources Fig. S8). In other words, for every unit of precipitation increase or decrease, discharge will change by a factor of 2. The elasticity curves are similar between cases 1 and 2 and across the different RRMs.

## Discussion

### Effect of potential evapotranspiration estimation on the robustness of rainfall-runoff models

The temporally varying PET (case 1) and long-term daily mean PET (case 3) produced similar RRM performance for the calibration period. This result is supported by several studies (Fowler 2002; Oudin et al. 2005b), although it is more accentuated in Northern Tunisia given the region's semi-arid climate condition and water-limited evaporation.

Our results illustrate a similarity between the different RRM in terms of sensitivity to PET information in spite of their different model structures, and their different formulations of calculating AET. This concept shows that long-term daily averaged PET does not affect RRM prediction capacity needed for hydrological projections. Although this result is supported by Oudin et al. (2005b), it does not match with those obtained by Guo et al. (2017), which showed that, for the case of several Australian catchments, the different evaporation process representations in conceptual rainfall-runoff models can have substantial impacts on discharge projections under a changing climate. One explanation for the disagreement between our results and that of Guo et al. (2017) is that their Australian catchments include a greater variety of climate (precipitation ranging between 344 and 1979 mm). Rather, under Northern Tunisia's semi-arid climate, the limited water availability heavily constrains AET.

### Sensitivity of hydrological projections to potential evapotranspiration estimation

The use of different PET estimations results in similar hydrological projections, similar elasticity curves, and also constitute a small proportion of the total variance within an ANOVA. This shows the low sensitivity of hydrological projections to PET estimation in Northern Tunisia. This means that it is relatively unimportant which PET formulation is used within hydrological climate change impact assessments for this region, and thus, a simple, temperature-based PET formel would be sufficient. To this extent, a bias in PET due to the use of such a formula would not have an important cumulative impact on the hydrological projections.

In arid regions, AET is highly limited by water availability, which limits the hydrological impact of any change in PET. The behaviour of AET, which is shown to increase with decreasing temperature (Fig. 3), could be explained by the negative correlation between precipitation and temperature in Northern Tunisia, e.g. dry years are associated with warm periods and wet years to cold periods. The opposite trends for the changes of AET and PET and the positively correlated change of AET and precipitation, under present and future climate conditions, indicate that precipitation is the main driver for AET in this region, and we confirm that the study catchments are water-limited rather than energy-limited.

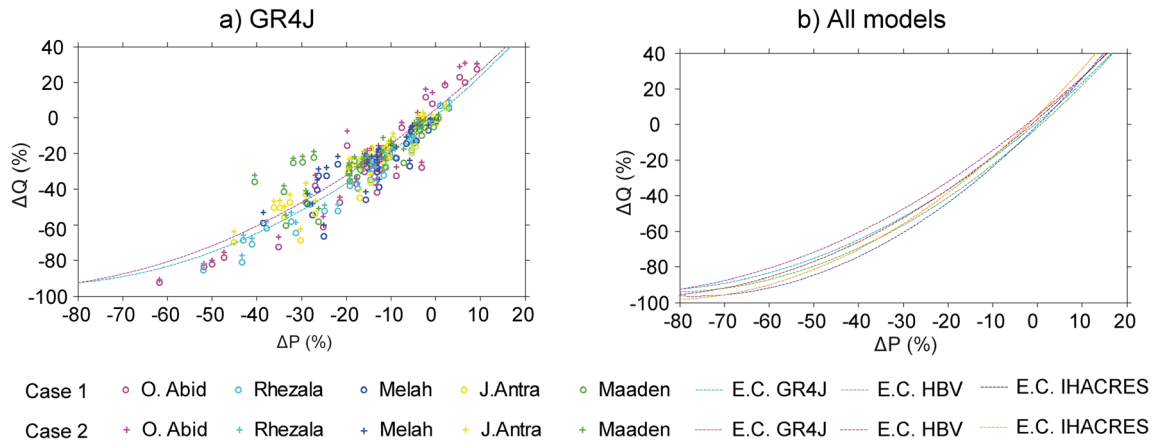



**Fig. 5** Elasticity of discharge due to precipitation change for (a) GR4J model and (b) shows elasticity curves (E.C.) from the other hydrological models. ΔP and ΔQ are the changes in discharge and precipitation from different rainfall-runoff models, catchments, representative concentration pathway scenarios (RCPs), climate models (GCM-RCMs), time periods and potential evapotranspiration estimation cases

### Impact of climate change on the hydroclimatology of Northern Tunisia

This study provides a first quantification on the impact of climate change on the discharge in Northern Tunisia. Our results show a significant future decrease in discharge due to a decrease in rainfall; this relationship is amplified because of the high elasticity of runoff to rainfall in this region. These results point to a situation of critical water stress in the future given that Tunisia already suffers from water paucity and is considered to be a water-poor country (Blinda and Thivet 2009). In this context, it seems that the competition between the ecological water requirements of wetlands with human uses will increase. The expected decrease of discharge, decrease of precipitation and increase of PET are factors that favour a future increase of salinity of the Lake Ichkeul. According to Hamdi et al. (2012), any change in the salinity pattern of the lake and its fresh water supply could affect negatively the available food of herbivorous waterbirds (mainly *Potamogeton* and *Scirpus* consumed by ducks and coots) and could affect the abundance of the waterbird communities. These results motivate adaption strategies for future climatic conditions, especially for the management of surface water resources and to ensure the ecological water requirement of wetland, a vital interest in the preservation of biodiversity.

### Limitations

We recognize that the projected PET, if calculated by a physically based formula such as Penman-Monteith, could be larger than the PET projected by a temperature-based formula, or lower than the present PET condition, which means that it is possible that cases 1 and 2 may not envelop all the possible realizations of PET. Evaporative demand depends mainly on four major physical variables: air temperature, vapour pressure, net radiation and wind speed. Temperature-based formulae consider only one of these variables: temperature. A substantial change in vapour pressure and/or increase of net radiation and/or increase of wind speed that accompanies the expected future increase in temperature could generate a possible increase in PET greater than that projected by temperature-based formula. Hence, any future increase of vapour pressure and others could compensate the increase in PET and could generate a decrease in PET compared with the present condition, as found for example in the pan evaporation paradox. According to Guo et al. (2017), who evaluated the sensitivity of PET to different climate drivers under different climate conditions in Australian catchments, PET is generally more sensitive to perturbations in temperature than to the other climate variables. In addition, they found that discharge generally shows a higher sensitivity to perturbations in temperature. In the case of the current study, while the temperature is the primary driver of PET, its change does not affect projected discharge to a large extent. We can therefore expect that the effect of perturbations in other climate variables would have a smaller impact compared with temperature. In addition, the future change in these variables could compensate each other. Wang et al. (2017) noticed for example that for several Chinese catchments, the decrease in relative humidity together with the negative contribution of wind speed are strong enough to neutralize the warming signal, leading to a strengthening in evaporative demand, over the period 1994–2014, in some regions.

## Conclusion

The methods proposed in this study were developed in order to simulate the impact of PET estimation in a data-scarce region, which could otherwise not support the data requirements for a full PET formulation comparison. We conclude that for Northern Tunisia, and for similar regions, a simple temperature-based PET formulation is sufficient for hydrological projections. The uncertainty introduced by the PET estimation is minimal compared with uncertainty stemming from RRM structure, RCP scenario and the GCM-RCM structure. While we acknowledge that the effect of PET estimation might be different for other regions and climatic regimes, these results confirm that the choice of any particular PET estimation method has a limited impact on discharge simulations for semi-arid catchments. This study shows that data scarcity in developing countries may not hinder creating reliable hydrological projections. This work may facilitate further investigation into the impact of climate change on Northern Tunisia's water resources, as well as for other data-scarce, arid regions.

**Acknowledgements** This work is part of the postdoctoral research of the first author at the University of Zurich, Switzerland, which was supported by a Swiss Government Excellence Scholarship. The authors are grateful to INM (Institut National de la Météorologie) and to DGRE (Direction Générale des Ressourcesen Eau) in Tunisia for providing the necessary hydro-climatic data for this study. We also recognize the EURO-CORDEX community for the provision of climate modelled data. We thank Dr. Urs Beyerle for his assistance with the retrieval of EURO-CORDEX data. We also thank Dr. Christopher Reyer and Dr. Christophe Randin for their valuable comments, which helped to improve the manuscript.

**Funding Information** Open access funding provided by Swedish University of Agricultural Sciences.

## References


Abbaris A, Dakhlaoui H, Thiria S, Baragaoui Z (2014) Variational data assimilation with the YAO platform for hydrological forecasting evolving water resources systems. In: Understanding, Predicting and Managing Water-Society Interactions, vol 364. IAHS Publ, pp 1–6. https://doi.org/10.5194/piahs-364-3-2014

Addor N, Rössler O, Köplin N, Weingartner R, Seibert J (2014) Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments. Water Resour Res 50(10):7541–7562. https://doi.org/10.1002/2014WR015549

Andréassian V, Perrin C, Michel C (2004) Impact of imperfect potential evapotranspiration knowledge on the efficiency and parameters of watershed models. J Hydrol 286:19–35. https://doi.org/10.1016/j.jhydrol.2003.09.030

Aouissi J, Benabdallah S, Chabaane Z, Cudennec C (2016) Evaluation of potential evapotranspiration assessment methods for hydrological modelling with SWAT—application in data-scarce rural Tunisia. Agric Water Manag 174. https://doi.org/10.1016/j.agwat.2016.03.004

Ben Fraj W, Elloumi M, Molle F (2019) The politics of interbasin transfers: socio-environmental impacts and actor strategies in Tunisia. Nat Res Forum 43:17–30. https://doi.org/10.1111/1477-8947.12165

Bae D-H, Jung I-W, Lettenmaier DP (2011) Hydrologic uncertainties in climate change from IPCC AR4 GCM simulations of the Chungju Basin. Korea J Hydrol 401(1–2):90–105. https://doi.org/10.1016/j.jhydrol.2011.02.012

Bargaoui Z, Dakhlaoui H, Houcine A (2008) Modélisation pluie-débit et classification hydroclimatique. Rev Sci Eau 21:233–245. https://doi.org/10.7202/018468ar

Blinda M, Thivet G (2009) Ressources et demandes en eau en Méditerranée : situation et perspectives. Sécheresse 20:9–16. https://doi.org/10.1684/sec.2009.0162

Brigode P, Oudin L, Perrin C (2013) Hydrological model parameter instability: a source of additional uncertainty in estimating the hydrological impacts of climate change? J Hydrol 476:410–425. https://doi.org/10.1016/j.jhydrol.2012.11.012

Chen J, Brissette FP, Chaumont D, Braun M (2013) Finding appropriate bias correction methods in downscaling precipitation for hydrologic impact studies over North America. Water Resour Res 49(7):4187–4205. https://doi.org/10.1002/wrcr.20331

Coron L, Andréassian V, Perrin C, Lerat J, Vaze J, Bourqui M, Hendrickx F (2012) Crash testing hydrological models in contrasted climate conditions: an experiment on 216 Australian catchments. Water Resour Res 48:W05552. https://doi.org/10.1029/2011WR011721

Cramer W, Guiot J, Fader M, Garrabou J, Gattuso JP, Iglesias A, Lange MA, Lionello P, Llasat MC, Paz S, Peñuelas J, Snoussi M, Toreti A, Tsimplis MN, Xoplaki E (2018) Climate change and interconnected risks to sustainable development in the Mediterranean. Nat Clim Chang 8:972–980. https://doi.org/10.1038/s41558-018-0299-2

Dakhlaoui H, Bargaoui Z, Bárdossy A (2009) Comparaison de trois méthodes d'usage de la technique des voisins les plus proches en vue d'amélioration de la performance de l'algorithme SCE-UA appliqué pour le calage du modèle pluie-débit HBV. In: Hydroinformatics in Hydrology, Hydrogeology and Water Resources, vol 331. IAHS Publ, pp 139–153

Dakhlaoui H, Bargaoui Z, Bárdossy A (2012) Toward a more efficient calibration schema for HBV rainfall-runoff model. J Hydrol 444–445:161–179. https://doi.org/10.1016/j.jhydrol.2012.04.015

Dakhlaoui H, Ruelland D, Tramblay Y, Bargaoui Z (2017) Evaluating robustness of conceptual rainfall-runoff models under climate variability in Northern Tunisia. J Hydrol 550:201–217. https://doi.org/10.1016/j.jhydrol.2017.04.032

Dakhlaoui H, Ruelland D, Trambaly Y (2019) A bootstrap-based differential split-sample test to assess the transferability of conceptual rainfall-runoff models under past and future climate variability. J Hydrol. https://doi.org/10.1016/j.jhydrol.2019.05.056

Duan QY, Sorooshian S, Gupta V (1992) Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resour Res 28:1015–1031. https://doi.org/10.1029/91WR02985

Fowler A (2002) Assessment of the validity of using mean potential evaporation in computations of the long-term soil water balance. J Hydrol 256:248–263. https://doi.org/10.1016/S0022-1694(01)00542-X

Fowler KJ, Peel MC, Western AW, Zhang L, Peterson TJ (2016) Simulating runoff under changing climatic conditions: revisiting an apparent deficiency of conceptual rainfall-runoff models. Water Resour Res 52(3):1820–1846. https://doi.org/10.1002/2015WR018068

Gupta HV, Kling H, Yilmaz KK, Martinez GF (2009) Decomposition of the mean squared error and NSE performance criteria implications for improving hydrological modelling. J Hydrol 377:80–91. https://doi.org/10.1016/j.jhydrol.2009.08.003

Guo D, Westra S, Maier HR (2017) Impact of evapotranspiration process representation on runoff projections from conceptual rainfall-runoff models. Water Resour Res 53:435–454. https://doi.org/10.1002/2016WR019627

Hakala K, Addor N, Teutschbein C, Vis M, Dakhlaoui H, Seibert J (2019) Hydrological climate change impact modeling. In: Maurice P (ed) Encyclopedia of water: science, technology, and society. https://doi.org/10.1002/9781119300762.wsts0062

Hamdi N, Touihri M, Charfi F (2012) Diagnostic Ecologique du parc National Ichkeul (Tunisie) après la construction des barrages; Cas des oiseaux d'eaux. Revue d'écologie 67(1):41–62

Hartmann G, Bárdossy A (2005) Investigation of the transferability of hydrological models and a method to improve model calibration. Adv Geosci 5:83–87. https://doi.org/10.5194/adgeo-5-83-2005

Hawkins E, Sutton R (2009) The potential to narrow uncertainty in regional climate predictions. Bull Am Meteorol Soc 90(8):1095–1107. https://doi.org/10.1175/2009BAMS2607.1

Henia L (2008) (directed by) Atlas de l'eau en Tunisie. Publication de la Faculté des Sciences Humaines et Sociales, Tunis

Institut Tunisien des Etudes Stratégiques (2014) Système Hydraulique de la Tunisie à l'horizon 2030. Tunisia 222 pp.

IPCC – Intergovernmental Panel on Climate Change (2013) Climate Change 2013: The physical science basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. In: Stocker TFD, Qin G-K, Plattner M, Tignor SK, Allen J, Boschung A, Nauels Y, Xia V, Bex, Midgley PM (eds) . Cambridge University Press, Cambridge 1535 pp.

Jabloun M, Sahli A (2008) Evaluation of FAO-56 methodology for estimating reference evapotranspiration using limited climatic data-application to Tunisia. Agric Water Manag 95:707–715. https://doi.org/10.1016/j.agwat.2008.01.009

Jakeman AI, Littlewood IG, Withehead PG (1990) Computation of the instantaneous unit hydrograph and identifiable component flows with application to two small upland catchments. J Hydrol 117:275–300. https://doi.org/10.1016/0022-1694(90)90097-H

Kingston DG, Todd MC, Taylor RG, Thompson JR, Arnellet NW (2009) Uncertainty in the estimation of potential evapotranspiration under climate change. Geophys Res Lett 36:L20403. https://doi.org/10.1029/2009GL040267

Klemeš V (1986) Operational testing of hydrological simulation models. Hydrol Sci J 31:13–24. https://doi.org/10.1080/02626668609491024

Li Z, Chen Y, Shen Y, Liu Y, Zhang S (2013) Analysis of changing pan evaporation in the arid region of Northwest China. Water Resour Res 2013(49):2205–2212. https://doi.org/10.1002/wrcr.20202

Lindström G, Johanson B, Gardelin MPM, Bergström S (1997) Development and test of the distributed HBV-96 hydrological model. J Hydrol 201:272–288. https://doi.org/10.1016/S0022-1694(97)00041-3

Mansour M, Hachicha M, Mougou A (2017) Trend analysis of potential evapotranspiration case of Chott-Meriem region (the Sahel of Tunisia). Int J Agric Innov Res 5:703–708. https://doi.org/10.1016/j.agwat.2015.12.004

Marchane A, Tramblay Y, Hanich L, Ruelland D, Jarlan L (2017) Climate change impacts on surface water resources in the Rheraya catchment (High-Atlas, Morocco). Hydrol Sci J. https://doi.org/10.1080/02626667.2017.1283042

Meinshausen M, Smith SJ, Calvin K et al (2011) The RCP greenhouse gas concentrations and their extensions from 1765 to 2300. Clim Chang:109–213. https://doi.org/10.1007/s10584-011-0156-z

Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models–part I: a discussion of principles. J Hydrol 10:282–290. https://doi.org/10.1016/0022-1694(70)90255-6

Oudin L, Michel C, Anctil F (2005a) Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 1-can rainfall-runoff models effectively handle detailed potential evapotranspiration inputs? J Hydrol 303:275–289. https://doi.org/10.1016/j.jhydrol.2004.08.025

Oudin L, Hervieu F, Michel C, Perrin C, Andréassian V, Anctil F, Loumagne C (2005b) Which potential evapotranspiration input for a lumped rainfall-runoff model? Part 2: towards a simple and efficient potential evapotranspiration model for rainfall-runoff modelling. J Hydrol 303:290–306. https://doi.org/10.1016/j.jhydrol.2004.08.026

Perrin C, Michel C, Andréassian V (2003) Improvement of a parsimonious model for streamflow simulation. J Hydrol 279:275–289. https://doi.org/10.1016/S0022-1694(03)00225-7

Prudhomme C, Williamson J (2013) Derivation of RCM-driven potential evapotranspiration for hydrological climate change impact analysis in Great Britain: a comparison of methods and associated uncertainty in future projections. Hydrol Earth Syst Sci 17:1365–1377. https://doi.org/10.5194/hess-17-1365-2013

Ruelland D, Hublart P, Tramblay Y (2015) Assessing uncertainties in climate change impacts on runoff in Western Mediterranean basins. In: Hydrologic non-stationarity and extrapolating models to predict the future, vol 371. IAHS Publ, pp 75–81. https://doi.org/10.5194/piahs-371-75-2015

Seibert J (2003) Reliability of model predictions outside calibration conditions. Nord Hydrol 34:477–492. https://doi.org/10.2166/nh.2003.028

Seiller G, Anctil F (2014) Climate change impacts on the hydrologic regime of a Canadian river: comparing uncertainties arising from climate natural variability and lumped hydrological model structures. Hydrol Earth Syst Sci 18:2033–2047. https://doi.org/10.5194/hess-18-2033-2014

Seiller G, Anctil F (2016) How do potential evapotranspiration formulas influence hydrological projections? Hydrol Sci J. https://doi.org/10.1080/02626667.2015.1100302

Sellami H, Benabdallah S, La Jeunesse I, Vanclooster M (2015) Quantifying hydrological responses of small Mediterranean catchments under climate change projections. Sci Total Environ 543:924–936. https://doi.org/10.1016/j.scitotenv.2015.07.006

Sheffield J, Wood EF, Roderick ML (2012) Little change in global drought over the past 60 years. Nature 491:435–438. https://doi.org/10.1038/nature11575

Sperna Weiland FC, Tisseuil C, Dürr HH, Vrac M, Beek LPHV (2012) Selecting the optimal method to calculate daily global reference potential evaporation from CFSR reanalysis data for application in a hydrological model study. Hydrol Earth Syst Sci 2012(16):983–1000. https://doi.org/10.5194/hess-16-983-2012

Terink W, Immerzeel WW, Droogers P (2013) Climate change projections of precipitation and reference evapotranspiration for the Middle East and Northern Africa until 2050. Int J Climatol 33:3055–3072. https://doi.org/10.1002/joc.3650

Teutschbein C, Seibert J (2012) Bias correction of regional climate model simulations for hydrological climate-change impact studies: review and evaluation of different methods. J Hydrol 456-457:11–29. https://doi.org/10.1016/j.jhydrol.2012.05.052

Tramblay Y, Ruelland D, Hanich L, Dakhlaoui H (2016) Hydrological impacts of climate change in north African countries. Sub-chapter 2.3.1, The Mediterranean region under climate change. A scientific update, IRD Éditions, 736 p

Tramblay Y, Jarlan L, Hanich L, Somot S (2017) Future scenarios of surface water resources availability in North African dams. Water Resour Manag 32:1291–1306. https://doi.org/10.1007/s11269-017-1870-8

Vormoor K, Heistermann M, Bronstert A, Lawrence D (2018) Hydrological model parameter (in)stability–"crash testing" the HBV model under contrasting flood seasonality conditions. HSJ 63:991–1007. https://doi.org/10.1080/02626667.2018.1466056

Wang T, Zhang J, Sun F, Liu W (2017) Pan evaporation paradox and evaporative demand from the past to the future over China: a review. WIREs Water 4:e1207. https://doi.org/10.1002/wat2.1207

Wilcox BP, Seyfried MS, Breshears DD, Stewart BA, Howell TA (2003) The water balance on rangelands. In: Encyclopedia of Water Science, vol 791–4. Marcel Dekker, New York

PAPER V

Hydrol. Earth Syst. Sci., 24, 3815–3833, 2020
https://doi.org/10.5194/hess-24-3815-2020


Hydrology and Earth System Sciences
Open Access
EGU

# Risks and opportunities for a Swiss hydroelectricity company in a changing climate

**Kirsti Hakala**[1,a], **Nans Addor**[2,b], **Thibault Gobbe**[3], **Johann Ruffieux**[3], **and Jan Seibert**[1,4]

[1]Department of Geography, University of Zürich, Zürich, 8057, Switzerland
[2]Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, NR4 7TJ, UK
[3]Energy Board, Groupe E SA, Granges-Paccot, 1763, Switzerland
[4]Department of Aquatic Sciences and Assessment, Swedish University of Agricultural Sciences, Uppsala, 750 07, Sweden
[a]now at: Department of Infrastructure Engineering, University of Melbourne, Victoria, 3010, Australia
[b]now at: College of Life and Environmental Sciences, University of Exeter, Exeter, EX4 4RJ, UK

**Correspondence:** Kirsti Hakala (kirsti.hakala@unimelb.edu.au)

Received: 12 September 2019 – Discussion started: 7 October 2019
Revised: 8 April 2020 – Accepted: 4 May 2020 – Published: 29 July 2020

**Abstract.** Anticipating and adapting to climate change impacts on water resources requires a detailed understanding of future hydroclimatic changes and of stakeholders' vulnerability to these changes. However, impact studies are often conducted at a spatial scale that is too coarse to capture the specificity of individual catchments, and, importantly, the changes they focus on are not necessarily the changes most critical to stakeholders. While recent studies have combined hydrological and electricity market modeling, they tend to aggregate all climate impacts by focusing solely on reservoir profitability. Here, we collaborated with Groupe E, a hydroelectricity company operating several reservoirs in the Swiss pre-Alps, and we co-produced hydroclimatic projections tailored to support the upcoming negotiations of their water concession renewal. We started by identifying the vulnerabilities of their activities to climate change; together, we then selected streamflow and electricity demand indices to characterize the associated risks and opportunities. We provided Groupe E with figures showing the projected impacts, which were refined over several meetings. The selected indices enabled us to assess a variety of impacts induced by changes in (i) the seasonal water volume distribution, (ii) low flows, (iii) high flows, and (iv) electricity demand. This enabled us to identify key opportunities (e.g., the future increase in reservoir inflow in winter, when electricity prices have historically been high) and risks (e.g., the expected increase in consecutive days of low flows in summer and fall which is likely to make it more difficult to meet residual flow requirements). We highlight that the hydrological opportunities and risks associated with reservoir management in a changing climate depend on a range of factors beyond those covered by traditional impact studies. This stakeholder-centered approach, which relies on identifying stakeholder's needs and using them to inform the production and visualization of impact projections, is transferable to other climate impact studies, in the field of water resources and beyond.

## 1 Introduction

Hydropower is the most widely used renewable energy resource across the globe (Schaefli, 2015). Given this global importance, there is a growing need to support the adaptation of hydropower facilities and operations to changes induced by climate change. This need is particularly strong in mountainous catchments, which are the major source of streamflow for hydropower production and are particularly sensitive to climate change (Schaefli et al., 2007; Zierl and Bugmann, 2005). Electricity companies across Switzerland are renewing and renegotiating their water concessions, transforming their existing infrastructure, and considering investments in new regions and sectors (Barry et al., 2015; SWV, 2012). However, in the vast majority of these cases, tailored analyses of climate change impacts are not used (Tonka, 2015).

To anticipate climate change impacts on hydropower production and to develop adaptation strategies, it is essential

to account for end-user vulnerabilities and hydroclimatic changes at the local scale (Schaefli, 2015). Currently, the majority of studies that perform a climate change impact analysis focus on either the effect of climate change on the seasonal cycle or on extreme events (Addor et al., 2014; Etter et al., 2017; Finger et al., 2012; FOEN, 2012; Hänggi and Weingartner, 2012; Köplin et al., 2014; Lopez et al., 2009; Vano et al., 2010) but rarely on a combination of both. Furthermore, until recently, changes in streamflow (water supply) were typically analyzed in isolation and were usually not combined with projections of future electricity demand (Gaudard et al., 2013). In recent studies (Anghileri et al., 2018; Gaudard et al., 2018b; Savelsberg et al., 2018), the modeling of the electricity market has been combined with hydrological simulations to project potential revenue under climate change. These studies contribute to bridging the gap between economists and hydrologists and account for the interconnected nature of water and electricity, which is fundamental for sustainable hydropower development. However, their focus is still on the seasonal cycle (see Savelsberg et al., 2018, for a detailed overview of recent research on the impact of climate change on hydropower). The focus on particular streamflow indices is often determined by what climate and hydrological modelers perceive as most adequate and relevant (an approach commonly referred to as "top-down"). However, this does not necessarily correspond to the needs of the stakeholders in charge of designing adaptation strategies. Top-down studies typically provide an overview of the impacts of climate change on hydrological resources; however, for stakeholders to assess the future profitability of their operations, more specific and local information is often needed (Vano et al., 2018). Given the potential consequences and costs associated with climate change impacts, it is essential to reduce the risk of maladaptation, which can result from misunderstanding end-users' vulnerabilities to climate change or from ill-designed projections (Broderick et al., 2019). Robust adaptation measures that provide benefits under a range of climate change scenarios are especially valuable, as they reduce the risk of maladaptation. Prioritizing stakeholder involvement early on enables them to expose their concerns regarding climate change and to establish which potential future changes should be assessed as priorities. This stakeholder-centered approach is often referred to as "bottom-up" (Wilby and Dessai, 2010; Addor et al., 2015).

Here, we present a case study relying on a stakeholder-centered approach for creating hydrological and climatological projections tailored to support climate change adaptation and water concession negotiations. We collaborated with a Swiss electricity company that manages and has shares in several hydropower reservoirs in Switzerland. This project started with meetings with representatives from the company, thereby involving them in the design of the study from the beginning. We relied on their expertise and asked them to identify which hydroclimatic changes their hydropower operations are most vulnerable to and to indicate change thresholds beyond which their activities would be significantly impacted. These meetings enabled us to pinpoint vulnerabilities of the company's operations to climate change and to select hydrological and electricity demand indices to characterize the associated risks. The representatives stated that they expect the following to be considered during concession negotiations (i) the development of the electricity market and competitors, (ii) the projected supply of water resources, (iii) the changes in electricity demand, and (iv) the costs associated with adhering to new environmental standards. This study focuses on the estimation of future water resources (point ii) and provides preliminary insights into future electricity demand (point iii). Hence, over the course of this study, we addressed the following research questions:

1. Climate change impacts on water resources are already broadly described by the scientific literature and in reports published by public entities (e.g., environmental agencies). While this broad-scale information is available to hydroelectricity companies, is it adequate to support their negotiations for concession renewal?

2. Future climate change impacts are uncertain and are typically communicated using an ensemble of simulations. How well do stakeholders incorporate this uncertainty into their decision-making process on adaption strategies?

3. Future reservoir profitability depends on a wide range of economic and environmental factors. How can projections focused on the availability of water resources be leveraged in the negotiation process of a concession, and what are their limitations?

This paper is organized as follows: Sect. 2 introduces the electricity company, the hydropower installations considered for this project, and describes the indices and associated thresholds selected by the electricity managers. Section 3 describes the observational and modeled data as well as the modeling framework employed to carry out hydrological (water supply) and climatological (electricity demand) projections. Section 4 presents the projected changes in the indices chosen by the electricity managers. Section 5 discusses the implications of these changes for future hydropower operations and possible future extensions of this study. In Sect. 6, we summarize our results and draw conclusions regarding the use of stakeholder-centered approaches in climate change impact analyses.

## 2 Project scope and identification of vulnerabilities to climate change

### 2.1 Hydropower company and study catchments

For this study, we interacted with two Groupe E electricity managers and helped them to assess future climate

change impacts. Groupe E is headquartered in Granges-Paccot in the canton of Fribourg in Switzerland. Considering all of Groupe E's installations and purchases from the electricity market, the company distributes an average of 3091 GWh $yr^{-1}$ to nearly 400 000 inhabitants and companies. The company's electricity generation fleet consists of 6 dams and 10 power stations. The installations are mainly located either directly along the Sarine River or on one of its tributaries (also on the Doubs, Wysswasser and the Binna River). Groupe E produces 1330 GWh of electricity yearly, which is approximately 43 % of the electricity that they distribute. The remaining 57 % is balanced by purchasing and trading on the electricity market.

This study focuses on the inflow into two of Groupe E's reservoirs: (i) the Vernex (Rossinière) dam – Montbovon power station and (ii) the Montsalvens dam – Broc power station (Fig. 1). The catchments of Montsalvens and Vernex dams have areas of 172.7 and 398.5 $km^2$, respectively (Table 1). The Vernex and Montsalvens installations are situated upstream of several other installations belonging to Groupe E, which turbinate water from the Sarine River along its lower reach. Given the placement of the Montsalvens and Vernex installations, their future functionality and security are crucial for Groupe E. We explored the future inflow into these two reservoirs in order to support adaptation to climate change and, in particular, the negotiation of a new water concession for the two installations, as discussed in Sect. 2.2. Groupe E is familiar with the ensembles and uncertainties associated with hydrological simulations, as they use ensembles of short-term hydrological forecasts for their daily operations. Groupe E was very transparent throughout this collaboration; however, Groupe E's future strategies are confidential and cannot be fully disclosed in this paper.

### 2.2 Negotiations of the water concessions

In Switzerland, the sovereignty of public waters is assigned to the cantonal or local/municipal authorities, which can grant the right to use water for electricity production to a hydropower company via a lease known as a concession (Mauch and Reynard, 2004). Most dams in Switzerland were built between 1945 and 1970, and water concessions were then typically granted for a maximum of 80 years. Therefore, many electricity managers are currently faced with challenges spurred on by the cessation of their water concessions (SWV, 2012). Lac du Vernex is a reservoir with concession agreements with the cantons of Vaud and Fribourg that are both due to end in 2052. Lac de Montsalvens is a reservoir located in the canton of Fribourg and has a concession agreement with the canton of Fribourg ending in the year 2076. Typically, the submission for renewal is due 15 years in advance (i.e., the submission for renewal is due in 2037 for Vernex and 2061 for Montsalvens). Given the liberalization of the Swiss electricity market, new competitors are entering previously closed markets. Therefore, some hydropower companies may consider the early renewal of their concessions (decades in advance) to ensure their production portfolio and to position themselves securely in the market. Projections of climate change on relevant streamflow indices offer electricity companies insight into their resource availability in the future, and they also help them gauge the flexibility of future operations.

During concession negotiations, the authorities granting the water rights and electricity managers will agree upon the duration of the contract and the terms of the water fee (i.e., the price to be paid by the electricity company to the owner of the water rights). The water fee is determined based on the gross capacity of the hydropower plant and elevation differential (head) as well as the amount of water that can be used for electricity production under particular hydrological conditions as defined in the concession (Betz et al., 2019). A key aspect in the negotiations of a water concession are new environmental regulations that hydropower companies must now comply with, such as new residual water flow requirements. Environmental impacts on the ecosystem were not a primary concern in the early stages of hydropower in Switzerland (Tonka, 2015). However, it is now well understood that hydropower systems impact the natural connectivity, temperature, and dynamics of rivers and, therefore, have substantial impacts on the downstream ecosystem (e.g., fish habitat). Swiss environmental regulations are listed within the Water Protection Act (Gewässerschutzgesetz), which sets the rules for residual water flow; it defines residual flow as the amount of water that must remain in a river after water withdrawals. Cantonal requirements are currently being strengthened to increase the amount of residual flow required to remain in streams, which reduces the amount of water for hydropower production (as discussed further in Sect. 5.1.2).

### 2.3 Vulnerabilities to climate change and the selection of indices and thresholds

Our discussions with Groupe E representatives enabled us to identify three main types of vulnerabilities: (i) water volume vulnerabilities (will seasonal changes in inflow distribution impact the reservoir profitability, given that electricity prices have historically been highest in winter as electricity demand is relatively higher during this season?), (ii) low-flow vulnerabilities (will low-flow situations become more frequent and make it more challenging to guarantee a residual discharge?), and (iii) high-flow vulnerabilities (will high-flow situations become more frequent and how may they be used for profit?). To address these vulnerabilities, streamflow indices were selected in collaboration with electricity managers. Corresponding thresholds were also chosen, whose exceedance would significantly impact Groupe E's production activities and profit. These hydrological indices and their relevance for hydropower operations are summarized in Table 2. While future changes in the mean monthly streamflow cycle have been well explored (Addor et al., 2014; Smi-

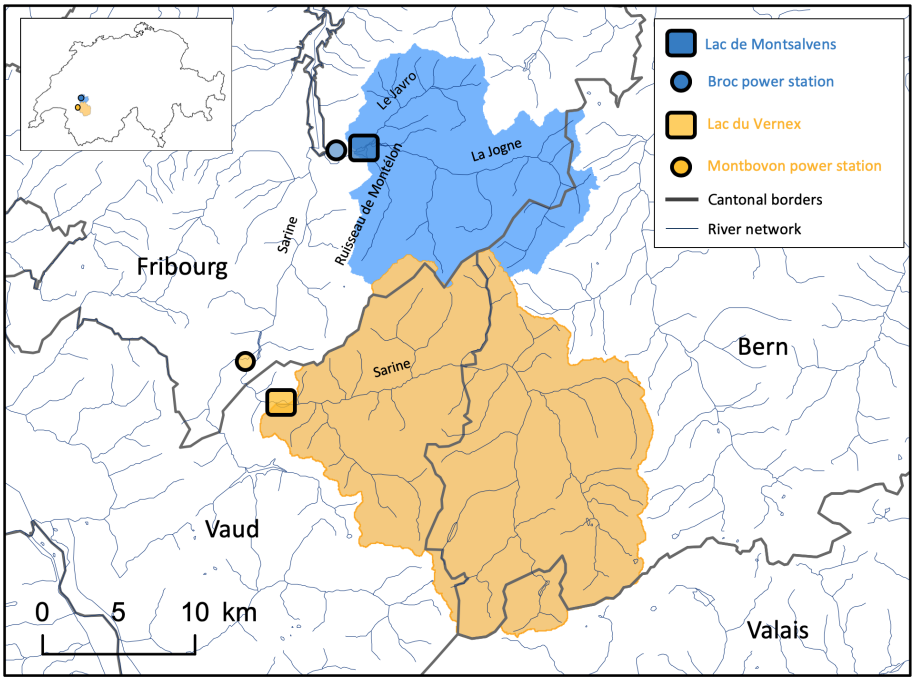



**Figure 1.** Map of the two study catchments: Montsalvens (blue) and Vernex (orange). The river network is shown in blue (dataset provided by the Swiss Federal Office for the Environment; FOEN); the cantons are labeled, and the dark gray lines depict the cantonal boundaries. The major river tributaries to the reservoirs are also labeled. The inset shows the location of the catchments within Switzerland.

**Table 1.** Main characteristics of the two study catchments, including catchment area, elevation, glacier coverage, karst percentage, forest cover, and energy production. Data for this table were derived from multiple sources: the area and mean elevation of the catchment were provided by Groupe E and were confirmed during delineation for modeling purposes, glacier coverage was estimated using satellite imagery from Google, karst hydrogeology was estimated using a dataset provided by Bitterli et al. (2004), and mean energy production was provided by the Swiss Federal Office of Energy (SFOE).

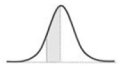
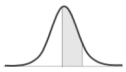
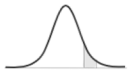
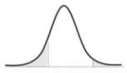
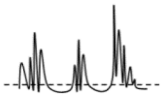
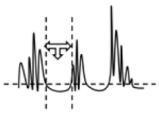
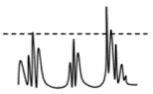
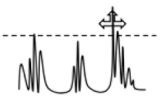
| Reservoir, dam | Area ($km^2$) | Mean elevation (m) | Glacier coverage (%) | Karst hydrogeology (%) | Mean energy production ($MWh\,yr^{-1}$) |
|---|---|---|---|---|---|
| Montsalvens, Broc | 172.7 | 1386 | 0 | 35 | 71 567 |
| Vernex, Montbovon | 398.5 | 1639 | < 1 | 15 | 59 422 |

atek et al., 2012; Vicuna and Dracup, 2007; Zierl and Bugmann, 2005), studies focusing on changes in other streamflow characteristics, such as extremes (Köplin et al., 2014), are less common. Groupe E representatives stated that although changes in the long-term mean monthly cycle are crucial, additional hydrological indices are necessary to inform their concession negotiations and adaptation efforts.

Aside from hydrological indices, Groupe E also requested an assessment of the rain versus snow contribution to runoff so that they can gain insight into their seasonal-scale operations. Historically, the Vernex and Montsalvens reservoirs reach their highest level in May after the spring runoff. The onset of the convective storm season is also around May/June. Thus, the coincidence of meltwater and high-intensity precipitation events can lead to excess storm flow entering the reservoirs; this excess water must be released without turbination, resulting in a profit loss and possible damage downstream. We used a hydrological model to characterize the respective contribution of rain and snowmelt to discharge (see Sect. 3.3.1).

Finally, two indices were chosen by Groupe E to gain insights into future electricity demand: cooling degree days ($C_{DD}$) and heating degree days ($H_{DD}$). They were computed following the method presented in Gaudard et al. (2013) and are solely based on air temperature as shown in Eqs. (1) and (2):

**Table 2.** Hydrological indices selected after discussions with Groupe E representatives. The relevance of each index for Groupe E's operations is explained, and the vulnerability thresholds for each index are provided. Relative changes exceeding these thresholds would have a significant impact on Groupe E's operations. In cases where two thresholds are provided, the exceedance of the lower threshold represents a significant impact and the upper threshold represents a critical impact. Visual aids for each index are also provided in the far-right column.

| Category | Hydrological index (season) | | Specific relevance for Groupe E | Vulnerability thresholds | Visual |
|---|---|---|---|---|---|
| Water volume | Long-term seasonal mean | March, April, May (MAM) | Snowmelt runoff may coincide with high-intensity precipitation events that could overwhelm Groupe E's ability to store and turbinate reservoir inflows. | 20%, 50% |  |
| | | June, July, August (JJA) | Meeting water demand for recreation, esthetic, and residual flow requirements. | |  |
| | | September, October, November (SON) | Managing reservoir level given drought concerns and meeting demand for recreation, esthetic, and residual flow requirements. | 20%, 50% |  |
| | | December, January, February (DJF) | Meeting energy demand during the coldest time of year and identifying opportunities to benefit from historically high energy prices. | |  |
| Low flows | Q5: 5th percentile of daily streamflow | June, July, August (JJA) | Meeting water demand for recreation, esthetic, and ecological purposes. Important for water fee negotiations and to assess whether regulations for residual flows are realistic. | 50% |  |
| | | September, October, November (SON) | Meeting water demand for recreation, esthetic, and ecological purposes. Important for water fee negotiations and to assess whether regulations for residual flows are realistic. | | |
| | Consecutive days of low flow | June, July, August (JJA) | Help with storage management during extended drought; concern for meeting energy demand and residual flow requirements. | 60 d |  |
| | | September, October, November (SON) | Storage management during extended drought, concern for meeting energy demand and residual flow requirements. | | |
| High flows | Q95: 95th percentile of daily streamflow | June, July, August (JJA) | Reservoir levels are at their highest, multi-day high-intensity precipitation can lead to water release without turbination (profit loss) or damage downstream in extreme cases. | 50% |  |
| | | December, January, February (DJF) | Explore opportunities to benefit from high-energy prices during the winter months, although DJF has not produced as much inflow as JJA. | | |
| | Consecutive days of high flow | June, July, August (JJA) | Reservoir level management and balancing dam releases with high intensity precipitation events. | 10 d |  |
| | | December, January, February (DJF) | Utilizing high volumes of inflow at times when the unit price of water is its highest. | | |

$$H_{\mathrm{DD}} = \max(\mathrm{Th} - \theta t,\, 0), \quad (1)$$

$$C_{\mathrm{DD}} = \max(\theta t - \mathrm{Tc}, 0) \quad (2)$$

Here, $\theta t$ is the air temperature retrieved from climate projections (Sects. 3.2.2 and 3.2.3). The thresholds Th = 13 °C and Tc = 18.3 °C were provided by Groupe E and correspond to the threshold values used in Gaudard et al. (2013). They represent the air temperatures that, when reached, cause consumers to turn on either cooling or heating in their homes. $C_{\mathrm{DD}}$ and $H_{\mathrm{DD}}$ were calculated for the cities (canton boundaries) of Zürich and Geneva, given that these areas are comprised of typical Groupe E electricity consumers. Results for Geneva are shown below, and results for Zürich can be found in Fig. S8 in the Supplement.

## 3 Data and methods for impact modeling

### 3.1 Modeling framework

To assess future changes in the streamflow and electricity demand indices introduced above, we relied on the following model chain. We combined 2 greenhouse gas emission scenarios (see Sect. 3.2.1), 11 regional climate models forced by general circulation models (GCM-RCMs; see Sect. 3.2.2), 2 GCM-RCM post-processing methods (see Sect. 3.2.4), and 1 hydrological model to simulate inflow entering the two reservoirs (Fig. 1). The hydrological model was calibrated using 3 objective functions, and 10 optimized parameter sets were generated per objective function and per calibration period (see Sect. 3.3.3). This modeling framework follows the procedure outlined in Hakala et al. (2020). It enabled us to assess uncertainties in the projected discharge and to provide Groupe E with a projected likely range for each index under future climate. The following subsections describe the steps of our modeling chain in greater detail.

### 3.2 Climate data and preparation

#### 3.2.1 Emission scenarios

Representative concentration pathways (RCPs) are scenarios describing possible futures for the evolution of Earth's atmospheric composition and, as such, provide boundary conditions for climate models. RCP4.5 and RCP8.5 were selected for this study. RCP4.5 corresponds to an intermediate emission trajectory, where greenhouse gas (GHG) emissions peak around 2040 and then generally stabilize. In contrast, RCP8.5 assumes that GHG emissions will continue to increase throughout the 21st century (Meinshausen et al., 2011).

#### 3.2.2 Observational and GCM-RCM data

Observational meteorological data were retrieved from the 2 km MeteoSwiss TabsD (Frei, 2014) and RhiresD (Frei and Schär, 1998; Schwarb, 2000) gridded datasets. The daily reservoir inflow was estimated by Groupe E for the period from 2008 to 2018 by solving the water balance based on variations of the reservoir level, the volume of water turbinated for hydropower production, and estimated losses due to evaporation from the reservoir (reservoir losses to the groundwater were neglected). GCM-RCM temperature and precipitation data were retrieved from the Coordinated Regional Downscaling Experiment for Europe (EURO-CORDEX; http://www.euro-cordex.net/, see Table 3). GCM-RCM model selection followed the methodology described in Hakala et al. (2018), which entails selecting models based on their hydrological performance over the historical period. Furthermore, we excluded models generating snow towers because of the influence that cooler temperatures associated with the snow towers may have on the climate change signal (Frei et al., 2018; Hakala et al., 2018; Zubler et al., 2016). EURO-CORDEX provides simulations at both 0.44° and 0.11° resolutions, but only 0.11° data were used given the size of the catchments investigated in this study. Overall, the exclusion of some GCM-RCMs due to their poor hydrological performance resulted in a tailored modeling setup that prioritized end-user decision-making.

#### 3.2.3 Data extraction

To extract temperature and precipitation from the gridded datasets, an area-weighted method, as shown in Hakala et al. (2018), was used. As a first step, the grid cells of the meteorological data were overlaid with the shapefile of a given catchment. Once the data from the overlapping grid cells were extracted, a weight factor was applied to each grid-cell time series based on the percentage of the catchment area overlapped by the grid cell, resulting in a single catchment-mean time series. This area-weighted methodology was used to extract temperature and precipitation data from both the EURO-CORDEX and MeteoSwiss datasets. In the case of the EURO-CORDEX dataset (horizontal grid spacing of $\sim$ 12.5 km), nine grid cells at least partially overlapped with the Vernex catchment and four grid cells overlapped with the Montsalvens catchment.

#### 3.2.4 Bias correction

The GCM-RCM simulated temperature ($T$) and precipitation ($P$) time series were bias corrected using a non-parametric quantile transformation of seasonal distributions. The cumulative distribution functions (CDFs) were determined individually for the different seasons – December–February (DJF), March–May (MAM), June–August (JJA), and September–November (SON) – for both the observed

**Table 3.** Overview of the 11 EURO-CORDEX GCM-RCM combinations used in this study. Some models were removed from the ensemble due to either snow tower issues or irregularities in the discharge simulations. The models that were removed are denoted using italic font.

| No. | GCM | RCM | Calendar | Notes |
|---|---|---|---|---|
| 1 | CNRM-CERFACS-CNRM-CM5 | CLMcom-CCLM4-8-17 | Gregorian | |
| 2 | ICHEC-EC-EARTH | CLMcom-CCLM4-8-17 | Gregorian | |
| 3 | MOHC-HadGEM2-ES | CLMcom-CCLM4-8-17 | 360 | |
| 4 | MPI-M-MPI-ESM-LR | CLMcom-CCLM4-8-17 | Gregorian | |
| | *ICHEC-EC-EARTH* | *DMI-HIRHAM5* | *Gregorian* | *R-ST* |
| | *NCC-NorESM1-M* | *DMI-HIRHAM5* | *Gregorian* | *R-ST* |
| | *IPSL-IPSL-CM5A-MR* | *IPSL-INERIS-WRF331F* | *Gregorian* | *R-D* |
| | *ICHEC-EC-EARTH* | *KNMI-RACMO22E* | *Gregorian* | *R-ST* |
| | *ICHEC-EC-EARTH* | *KNMI-RACMO22E* | *360* | *R-ST* |
| 5 | MOHC-HadGEM2-ES | KNMI-RACMO22E | Gregorian | |
| 6 | MPI-M-MPI-ESM-LR | MPI-CSC-REMO2009 | Gregorian | |
| 7 | CNRM-CERFACS-CNRM-CM5 | SMHI-RCA4 | Gregorian | |
| 8 | ICHEC-EC-EARTH | SMHI-RCA4 | Gregorian | |
| 9 | IPSL-IPSL-CM5A-MR | SMHI-RCA4 | Non-leap year | C |
| 10 | MOHC-HadGEM2-ES | SMHI-RCA4 | 360 | |
| 11 | MPI-M-MPI-ESM-LR | SMHI-RCA4 | Gregorian | |

(R-ST) refers to models removed due to snow towers in the GCM-RCM model output.
(R-D) refers to models removed due to irregularities in the mean monthly distribution of discharge when simulations were forced with this GCM-RCM.
(C) denotes that the calendar was converted from non-leap year to proleptic Gregorian.

(MeteoSwiss) and simulated (EURO-CORDEX) $T$ and $P$ time series. For GCM-RCMs with a non-leap-year calendar (Table 3), $T$ and $P$ were converted to a Gregorian calendar prior to bias correction. For GCM-RCMs with a 360 d calendar, observational data were converted to a 360 d calendar before bias correction, and the hydrological model was run using this calendar. The "qmap" package in R (Gudmundsson, 2016; Gudmundsson et al., 2012) was used to match the CDF of the simulated data to that of the observed data. Specifically, a transfer function was generated to match each raw GCM-RCM $P$ and $T$ percentile to the associated $P$ and $T$ percentile of the MeteoSwiss data. The biases in the raw GCM-RCM simulations were assumed to be stationary over time; thus, the same transfer functions were used to correct the projections of $T$ and $P$.

### 3.3 Hydrological data and model

#### 3.3.1 Hydrological model

The bucket-type Hydrologiska Byråns Vattenbalansavdelning (HBV) model (Bergström, 1976; Lindström et al., 1997) was used to simulate streamflow entering the two reservoirs. For this project, we used the HBV-Light version (Seibert and Vis, 2012). HBV is a semi-distributed model that uses four routines (snow, soil, response, and routing routines) and relies on elevation bands to account for changes in $T$ and $P$ with elevation within a catchment. HBV requires temperature, precipitation, and potential evaporation time series as input. For a more detailed description of the separate routines, we refer the reader to Seibert and Vis (2012). For the remainder of the paper, we use the term HBV when referring to the HBV-Light version.

#### 3.3.2 Adjustment of discharge data

When initially analyzing the discharge data provided by Groupe E in combination with MeteoSwiss observational meteorological data, we noticed that precipitation was too small to explain the discharge flowing into the Montsalvens reservoir. Based on water balance calculations informed by karst hydrogeological information (Bitterli et al., 2004) and actual evaporation estimates (Menzel et al., 1999), it was assumed that karst was responsible for the larger than expected discharge. The Montsalvens and Vernex catchments are located in a transitional region between the Alps and the Swiss Plateau. As pointed out by Fan, (2019), a catchment is more likely to be an open or "leaky" system when positioned at either the high or low end of a steep regional topographic and climate gradient, which is the case here. Therefore, a correction factor was applied to the observed discharge to rescale it to match the expected mean discharge. The factor was calculated following the water balance equation $P = E + (f \cdot Q) + \Delta S$ for the period from 2008 to 2018, where $P$ represents precipitation falling within the catchment; $E$ stands for actual evaporation; $Q$ represents the inflow reported to enter the Montsalvens reservoir; and $\Delta S$ stands for change in storage, which was considered negligible in this case. By applying the factor $f$ (0.79) to the discharge time series, we were able to close the water balance equation. Therefore, this method assumes that 21 % of the total inflow entering the Montsalvens reservoir is groundwater

entering through the karst system. Karst hydrogeology did not appear to have a discernible effect on discharge for the Vernex catchment.

#### 3.3.3 Calibration and validation

Calibration and validation of HBV were based on three different objective functions, namely the Lindström measure (Lindström et al., 1997), the Nash–Sutcliffe efficiency (NSE; Nash and Sutcliffe, 1970), and the Kling–Gupta efficiency (Gupta et al., 2009). Two separate periods were used for calibration and validation: 1 October 2008 to 30 September 2013 and 1 October 2013 to 31 August 2018. For each combination of objective function and time period, 10 independent parameter sets were generated. HBV was calibrated using a genetic algorithm and Powell optimization (Seibert, 2000) method (10 000 model runs for the genetic algorithm and an additional 1000 runs for the Powell optimization). Using multiple objective functions and calibration periods enabled us to account for parameter uncertainty and to generate an ensemble of equally likely realities (Brigode et al., 2013; Coron et al., 2012; Klemeš, 1986). Both catchments achieved reasonable calibration and validation scores (an NSE of 0.75 or higher for all objective functions and periods). Therefore, all parameter sets were carried forward in the modeling chain.

### 3.4 Evaluation of the modeling chain over the reference period

Prior to creating projections, we analyzed our modeling chain performance over a reference period. Figure 2 provides a comparison between $(\text{variable})_{\text{obs}}$ and $(\text{variable})_{\text{ref}}$ for each hydrological index and climate change impact index. The ref subscript indicates that the index was computed using HBV simulations driven by observed atmospheric forcing. In the case of the hydrological indices, $Q_{\text{obs}}$ and $Q_{\text{ref}}$ stem from different time periods, as Group E records only cover the period from 2008 to 2018. Given this mismatch in time periods, we began by comparing the monthly precipitation of the $Q_{\text{obs}}$ and $Q_{\text{ref}}$ time periods (Supplement Fig. S1). The period from 1980 to 2009 ($Q_{\text{ref}}$ period) experienced a wetter climate than the period from 2008 to 2018 ($Q_{\text{obs}}$ period).

Figure 2 shows that hydrological simulations driven by raw climate simulations present severe biases. For instance, the mean monthly inflow is vastly overestimated by raw data from April through December (Fig. 2a). Bias correction leads to a significant reduction in these biases, and it was necessary to capture the indices required by Group E (Fig. 2a–h). Figure 2g shows that the application of bias correction is successful at reducing the ensemble spread of $H_{\text{DD raw}}$ (yellow shaded area), resulting in $H_{\text{DD qm}}$ (purple shaded area). $H_{\text{DD qm}}$ can be seen to fit well with $H_{\text{DD ref}}$ for the entirety of the annual cycle. Figure 2h also shows a reduction in the $C_{\text{DD raw}}$ ensemble spread (yellow shaded area) due to the application of quantile mapping ($C_{\text{DD qm}}$; purple shaded area), with August retaining a relatively high level of uncertainty. As concession negotiations require more finely tuned projections than what can be delivered by raw simulations, we excluded simulations generated using raw GCM-RCM data from the results section so that the focus can be on future changes and not on the effects of the bias correction. Figures displaying hydrological variables utilize two $y$ axes where specific discharge ($\text{mm d}^{-1}$) is shown on the left-hand axis, and discharge ($\text{m}^3\,\text{d}^{-1}$) is displayed on the right-hand axis. The former allows for a comparison between catchments, whereas the latter is more useful for electricity managers when operations are primarily looked at in terms of volumes.

Overall, when using bias-corrected climate simulations, HBV satisfactorily captures the annual discharge cycle (Fig. 2a), the respective contribution of snow and rain to streamflow (Fig. 2b), and Q5 and Q95 during the seasons of interest (Fig. 2c, d). In contrast, HBV tends to overestimate both the duration of periods below Q5 and above Q95 (Fig. 2e, f). It is, however, important to note that HBV was not specifically calibrated against the hydrological indices mentioned in Table 1; thus, it is not surprising if $Q_{\text{obs}}$ and $Q_{\text{ref}}$ deviate when compared across these indices.

### 3.5 Projections of climate change impacts

As the performance of the modeling chain was considered to be satisfactory over the reference period, all parameter sets generated in Sect. 3.3.3 were used to simulate projections for the periods from 2020 to 2049, from 2045 to 2074, and from 2070 to 2099. Our modeling chain was comprised of 2 emission scenarios (RCP4.5 and RCP8.5), 11 EURO-CORDEX GCM-RCMs, 2 post-processing methods (raw and quantile mapping), 1 hydrological model (HBV), 3 objective functions for the hydrological model (Lindström measure, Nash–Sutcliffe efficiency, and Kling–Gupta efficiency), and 10 optimized parameter sets per objective function and 2 calibration periods. This led to a total of 1320 bias-corrected simulations for each future period and basin. Below, we focus on the comparison between 1980–2009 and 2070–2099 under RCP8.5 and on the Vernex catchment. The results and figures for all periods, RCP4.5, and both catchments were provided to Groupe E, and the end-of-century results for Montsalvens can be found in the Supplement. The projected streamflow indices were not compared to observed discharge data, because such a comparison could be misleading due to the mismatch in time periods and the inclusion of hydrological model uncertainty. Instead, the projections were compared to simulations for the reference period based on bias-corrected GCM-RCM simulations.

## 4 Results

This section presents the changes in streamflow and electricity demand indices projected by our modeling chain. The im-

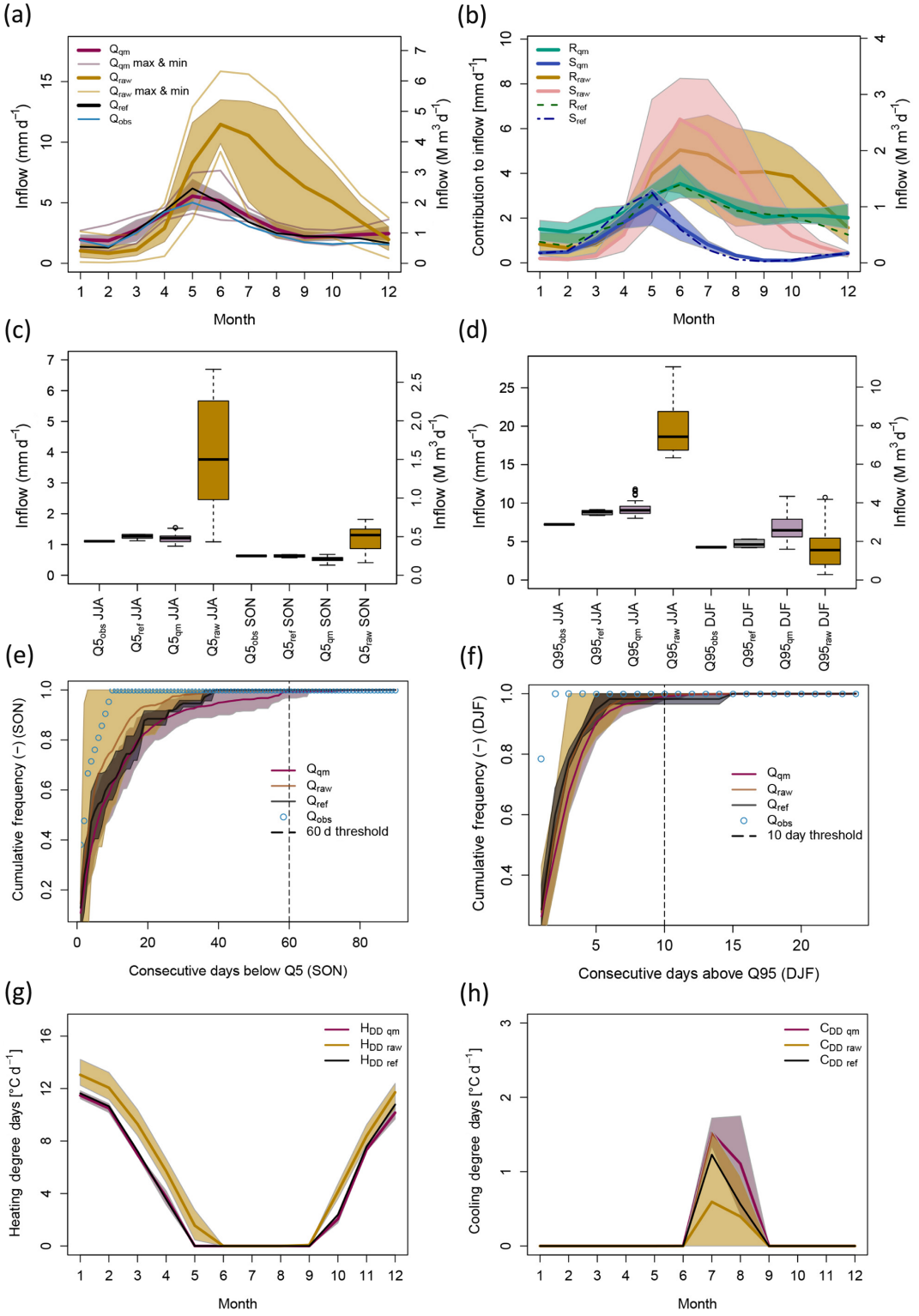



**Figure 2.** The performance of the calibration of HBV and the bias correction treatment are shown for each index for the Vernex catchment **(a–f)** and the canton of Geneva **(g, h)**. When observational data were not available, only the bias correction performance is shown **(g, h)**. All simulated data cover the period from 1 January 1980 to 31 December 2009, except for $Q_{obs}$ data which span the period from 1 October 2008 to 31 August 2018. Panels **(a)**, **(b)**, **(g)**, and **(h)** depict long-term monthly means.

plications of these changes for future reservoir operations and profitability are discussed in Sect. 5. ensemble lying close to the

### 4.1 Projected changes in water volume

Figure 3a compares the historical (1980–2009) and future (2070–2099) annual distribution of inflow entering the Vernex reservoir for RCP8.5. Changes in winter (DJF) discharge are shown to widely exceed the +20 % and +50 % thresholds specified by Groupe E. Meanwhile, the summer (JJA) discharge decrease is expected to be around the −50 % threshold (ensemble mean). Groupe E asked for the long-term mean monthly discharge cycle to be visualized by showing the volume difference between future (2070–2099) and historical (1980–2009) conditions. Figure 3b was requested so that the total amount of water gained/lost can be directly considered during concession negotiations. Under RCP8.5, the Vernex reservoir should experience more inflow between December and March but less inflow from May to October. By the end of the century, the expected average change in inflow for the Vernex reservoir is $-1.11\,\mathrm{M\,m^3\,d^{-1}}$ (likely range from −4.52 to +2.54) under RCP8.5 and $-0.24\,\mathrm{M\,m^3\,d^{-1}}$ (likely range from −2.97 to +2.35) under RCP4.5. Similarly, the inflow entering the Montsalvens reservoir is expected to experience an average decrease of $-0.72\,\mathrm{M\,m^3\,d^{-1}}$ (likely range from −2.19 to +0.81) under RCP8.5 and $-0.18\,\mathrm{M\,m^3\,d^{-1}}$ (likely range from −1.61 to +1.08) under RCP4.5.

The shift in the annual distribution of inflow entering the reservoirs is primarily caused by changes in the form of precipitation contributing to inflow (Fig. 4). The peak annual contribution to inflow from snowpack is expected to decrease by more than half and to occur earlier in the year, shifting from May to April. Spring runoff derived from snowpack will likely be a less reliable source of inflow in the future. Meanwhile, rain is shown to decrease its respective contribution to inflow over the summer. The shift in spring runoff and the reduction in the rainfall contribution to inflow results in a reduction in inflow entering the reservoirs (Figs. 3b, S4). Over the 21st century, winters are expected to see an increasing rain contribution to inflow and a reduced contribution from both rain and snow from May until November. The Montsalvens catchment is expected to experience a similar regime change in the future, with an even more pronounced reduction in snowfall contribution (Fig. S5).

### 4.2 Projected changes in low flows

$Q_{\mathrm{qm}}$ simulations of low flows (Q5) for JJA and SON strongly decrease under RCP8.5, with the majority of the ensemble indicating a decrease greater than the −50 % threshold (Fig. 5a). The spread of the ensemble for both seasons is relatively small in absolute terms. Projections for the inflow entering the Montsalvens reservoir indicate similar changes, with Q5 dropping below the −50 % threshold for JJA and the median of the SON ensemble lying close to the −50 % threshold (Fig. S6).

The frequency of consecutive days below Q5 is expected to increase under the influence of climate change in SON. Figure 6a demonstrates this concept by showing the cumulative distribution functions (CDFs) of the consecutive days below Q5 for the Vernex catchment over the SON season. The robust nature of the change compared with historical simulations demonstrates that there is high confidence that there will be more days below Q5 over the SON season in the future, although it should be noted that $Q_{\mathrm{qm}}$ data initially overestimated the CDFs of consecutive days below Q5 (Fig. 2e). The results for the Montsalvens reservoir agree with the changes shown for the Vernex reservoir, with a slightly less pronounced difference between the historical and future periods. For the Montsalvens catchment, there are relatively fewer extended periods of low flow (Fig. S7).

### 4.3 Projected changes in high flows

The magnitude of high flows (Q95) is expected to decrease in JJA under RCP8.5 (Fig. 5b). However, the median and the majority of ensemble members are within the 50 % threshold interval. In contrast, for winter, $Q_{\mathrm{qm}}$ simulations show a significant increase, far exceeding the +50 % threshold. Inflows entering the Montsalvens reservoir exhibit similar behavior over both seasons (Fig. S6).

More extended periods of consecutive days above Q95 are projected in DJF under the influence of climate change. The CDFs of the future simulations show a significant increase in the length of consecutive high-flow periods, including periods longer than the stipulated 10 d threshold. Results for Montsalvens indicate similar but less pronounced changes (Fig. S7).

### 4.4 Projected changes in temperature-based indices

Figure 7a shows that the number of $H_{\mathrm{DD}}$ is expected to decrease over the winter months under the influence of climate change, whereas the summer months experience no change as this time of year is already too hot to invoke heating within a household. Figure 7b shows that $C_{\mathrm{DD}}$ will likely increase for the months between May and October. The winter months show no change as these months are too cold to invoke cooling within the household of a typical electricity customer. Projections for the canton of Zürich show a general agreement with the magnitude and distribution of change (Fig. S8).

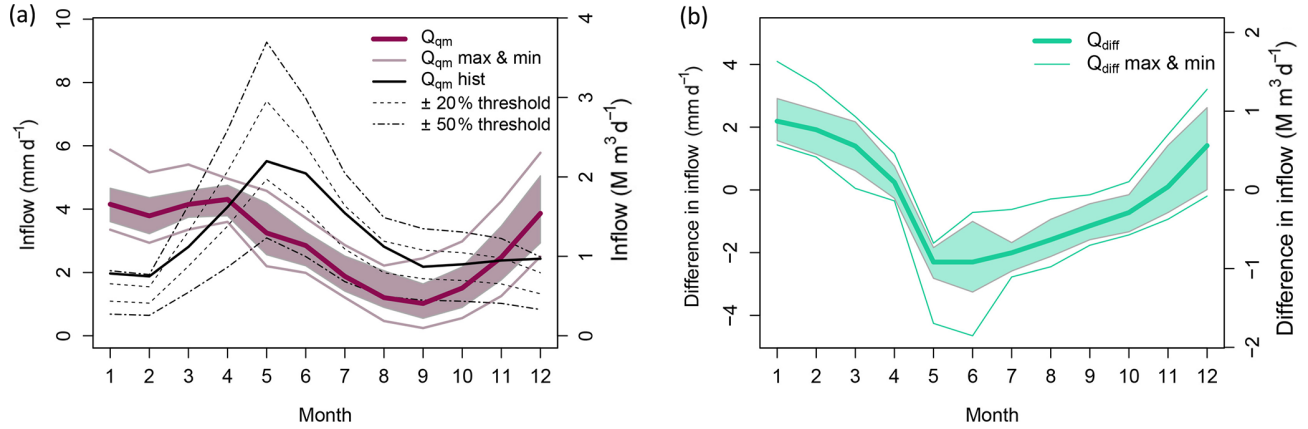



**Figure 3. (a)** Long-term mean monthly inflow entering the Vernex reservoir for 1980–2009 ($Q_{qm}$ hist) and for 2070–2099 ($Q_{qm}$, RCP8.5). The mean (solid lines) and likely range (shaded areas) are shown, where the likely range represents two-thirds of the simulations. The two thresholds are based on the mean of the simulations forced by observed climate data ($Q_{ref}$ over the period from 1980 to 2009). **(b)** Long-term mean monthly change in inflow (2070–2099 with respect to 1980–2009) for the Vernex catchment.

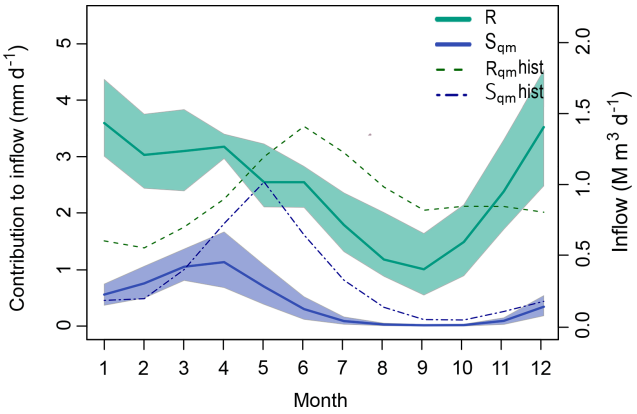



**Figure 4.** Mean monthly contribution of rain ($R$, green) versus snow ($S$, blue) to inflow entering the Vernex reservoir. Two periods are compared: 1980–2009 ($R_{qm}$ and $S_{qm}$ hist) and 2070–2099 ($R_{qm}$ and $S_{qm}$). All projections shown are simulations under RCP8.5. The mean (solid lines) and likely range (shaded areas) are shown, where the likely range represents two-thirds of the simulations. The dashed lines indicate the mean of the reference simulations.

## 5 Discussion

### 5.1 Implications of the projected changes for hydropower operations

The projected changes in streamflow are summarized in Fig. 8 along with the critical thresholds selected by Groupe E. Here, we discuss the implications of these changes for hydropower operations and how they can be used as leverage during the negotiation of the water concession.

#### 5.1.1 Water volume

Some changes in the seasonal inflow distribution represent new opportunities. Over the winter period, the inflow into Lac du Vernex is expected to increase by 90 % under RCP8.5 (Fig. 8a, b) and by 63 % under RCP4.5 (ensemble mean). Inflow into Lac de Montsalvens is expected to increase by 89 % under RCP8.5 and by 61 % under RCP4.5 (ensemble mean). Hydropower has the potential to remain an important source of electricity in the winter given the low yield of photovoltaics during the short winter days and the unpredictability and contentious politics of wind power (Kienast et al., 2017). Therefore, these changes could allow Groupe E to capitalize on generally higher electricity prices in winter (assuming that electricity prices remain higher in winter than in summer), resulting in a potential increase in profits for this season.

In contrast, regime changes in the summer and fall are expected to lead to new challenges for Groupe E. Over the summer period, Lac du Vernex is expected to experience an average decrease of −51 % under RCP8.5 (Fig. 8a, b) and −30 % under RCP4.5 (ensemble mean); Lac de Montsalvens is likely to experience an average decrease of −49 % under RCP8.5 and −28 % under RCP4.5. The reduction in summer inflow can be linked to the snowpack shrinkage over the coming century and the simultaneous reduction in total precipitation over the summer months (Fig. 4). Köplin et al. (2014) showed that when snow accumulation is important to a catchment hydrological regime during the historical period, the anticipated changes in seasonality are most pronounced. Groupe E stated that the Vernex and Montsalvens reservoirs are too small to store water over the winter period in order to offset droughts in the summer period. Adjusting the size of their reservoirs is currently not a viable option; therefore, it was not explored by our modeling experiments.

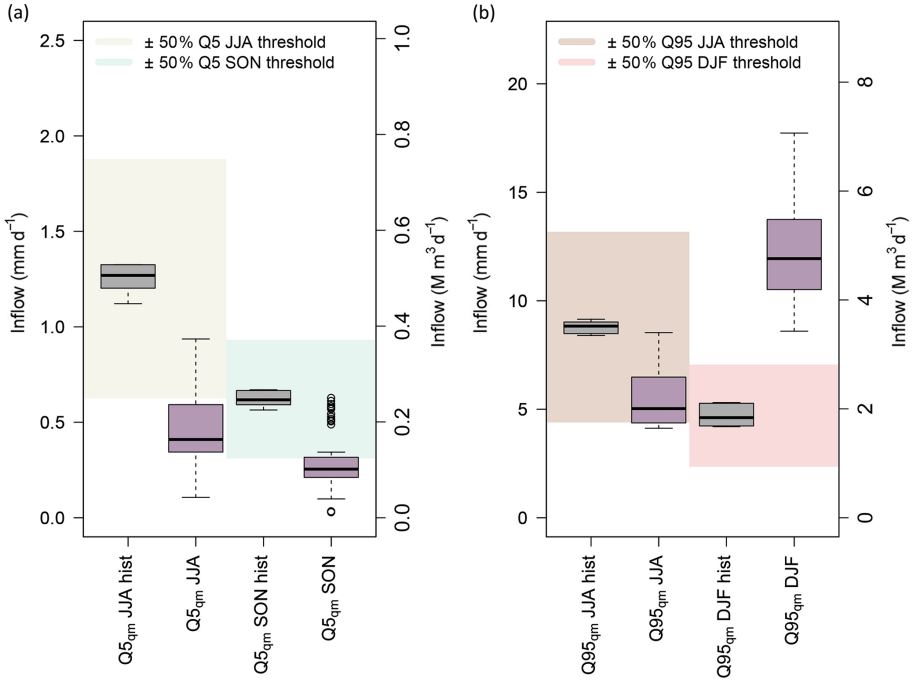



**Figure 5. (a)** Boxplots showing low-flow (Q5) and **(b)** high-flow (Q95) indices, where the historical period (gray boxplots; 1980–2009) is compared to the future period (purple boxes; 2070–2099) for inflows entering the Vernex reservoir. All projections shown are for RCP8.5. For each index, an associated ±50 % threshold is designated by the shaded area. These thresholds are based on the mean of simulations when forced by observed climate data ($Q_{\mathrm{ref}}$) over the period from 1980 to 2009.

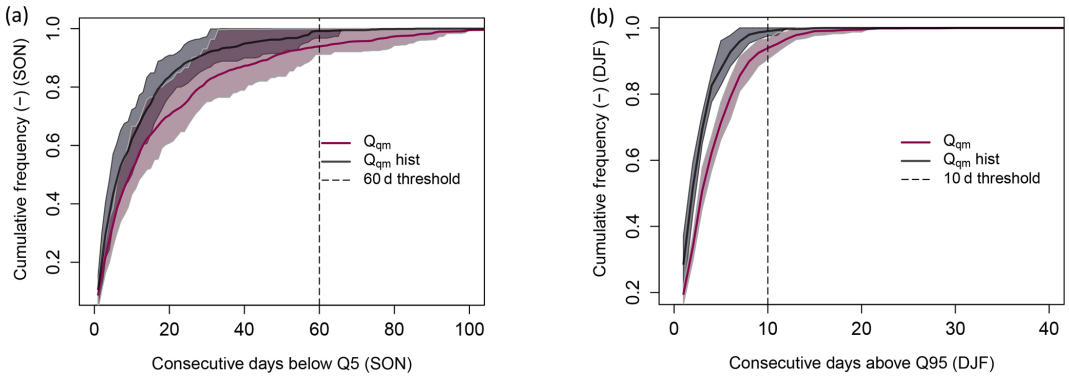



**Figure 6.** Cumulative distribution functions (CDFs) are shown, where the historical period (1980–2009; gray) is compared to the future period (2070–2099; purple) for the Vernex catchment. **(a)** CDFs for consecutive days below Q5 are shown for the SON season, and a 60 d threshold is indicated by the black dashed line. **(b)** CDFs of the consecutive days above Q95 are shown for the season of DJF, and a 10 d threshold is shown by the black dashed line. Instances where the simulations exceed their associated threshold represent a level of change that is of interest to Groupe E. The mean (solid lines) and likely range (shaded areas) are shown, where the likely range represents two-thirds of the simulations.

Given a decrease in inflow over the summer and a possible increase in electricity demand for cooling (Fig. 7b), an investment in other energy sources may be considered, such as photovoltaics which have their peak production during the longer summer days. In addition to other market conditions and legal requirements, hydropower energy providers may use these projections of changes in water volume to negotiate a lower cost for their water fee, as the fee is partially determined based on the amount of water that can be used for electricity production. An impact comparison of the different water fee systems on Swiss hydropower was performed by Gaudard et al. (2018a). Within their study, they compared different water fee frameworks including a (i) no-fee system, (ii) a fixed-fee system, (iii) a semiflexible or fixed and variable fee system, and a (iv) profit-based imposition system. The current water fee framework follows a fixed-fee system.

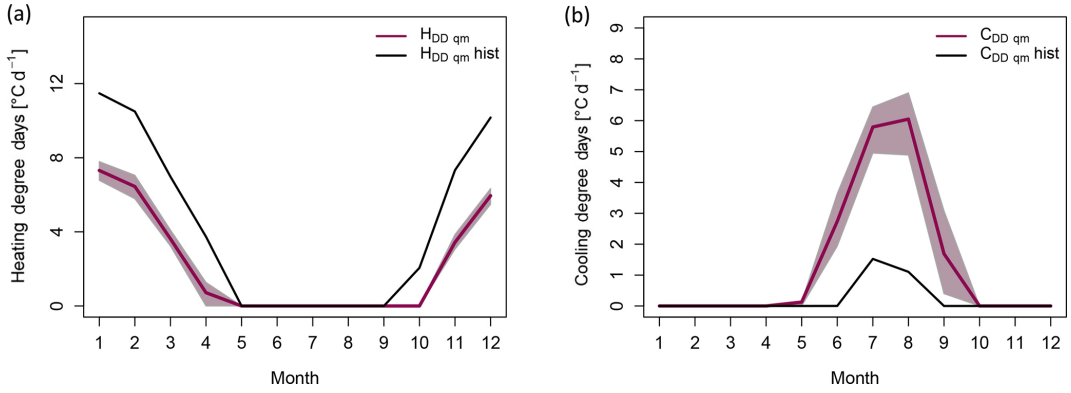



**Figure 7.** Mean monthly **(a)** $H_{DD}$ and **(b)** $C_{DD}$ for the canton of Geneva. The mean of the historical simulations (1980–2009; gray) are compared to the future simulations under the influence of RCP8.5 climate change scenario (2070–2099; purple). The mean (solid lines) and likely range (shaded areas) are shown, where the likely range represents two-thirds of the simulations. Groupe E prescribed thresholds of 13 and 18.3 °C to compute $H_{DD}$ and $C_{DD}$, respectively.

The authors discuss that the water fee tends to flatten the differences between the lowest and highest financial years under a fixed system. The hydropower sector is vulnerable in such a system, which provides no flexibility and instead imposes a fee based on theoretical power. In contrast, a profit-based system is shown to increase the financial robustness of the hydropower sector. The water fee framework is subject to review in 2024 (Betz et al., 2019).

#### 5.1.2 Low flows

Low flows will require special attention in the coming decades, as the magnitude of Q5 is likely to reduce drastically, with the majority of ensemble members predicting a change exceeding the −50 % threshold in JJA and SON (Fig. 8c, d). In addition, periods of low flow are expected to increasingly extend beyond Groupe E's 60 d threshold in JJA and SON (Fig. 8c, d). These changes are likely to influence the negotiated terms of the water fee. The decrease in production over a long period of time has a significant effect on the flexibility of production. Flexibility is a significant component of a storage hydropower plant's profitability, as it enables hydropower operators to turbinate when electricity prices are optimal.

Cantonal requirements are currently being strengthened to reduce environmental impacts. One of the cantonal measures includes increasing the amount of residual flow for environmental reasons (e.g., flora, fauna, and sediment transport are affected by very low flows). This study shows that the water carried by low flows is expected to substantially decrease over the coming decades, and the duration of low-flow conditions will likely increase. Hence, minimum flow requirements are likely to be a delicate topic during concession negotiations, as Groupe E may request that residual flow requirements do not increase, which is likely to be challenged by stakeholders that are primarily concerned with environmental issues.

#### 5.1.3 High flows

Opportunities are present over the winter period, as the average high inflows to the Vernex and Montsalvens reservoirs are projected to increase by more than 50 % (Fig. 8d) and exceed the 10 d threshold (Fig. 6b). An increase in high flows entering the reservoirs during the winter period, when electricity prices are highest, would allow Groupe E to better satisfy demand using their own production, rather than supplementing their supply by trading/purchasing on the electricity market. The hydrological shift from slow, snow-dominated processes to more variable, rainfall-driven processes will require a flexible operating framework so that these quick inflows can lead to increased profit, rather than spillover. Storage power plants are already being utilized to their full extent during peak price hours, so additional inflows in winter and early spring will be utilized in hours of lower prices (Savelsberg et al., 2018). To generate more revenue, the extra inflow would have to be captured and turbinated at optimal times or at prearranged prices. Groupe E could consider investing in their existing short-term forecasting and trading unit in order to improve their forecasts of high-flow events. As Groupe E can decide when to sell its electricity (anytime between the next hour to the next 3 years), a balance between best price and risk management needs to be found. Conversely, projections show a decrease in high flows in the summer (Fig. 8c), which indicates a reduced risk of water loss due to spillover events.

#### 5.1.4 Electricity demand

To adapt to climate change, hydroelectricity companies cannot base their strategies on water availability alone, they also need to estimate future electricity demand (Gaudard et al.,

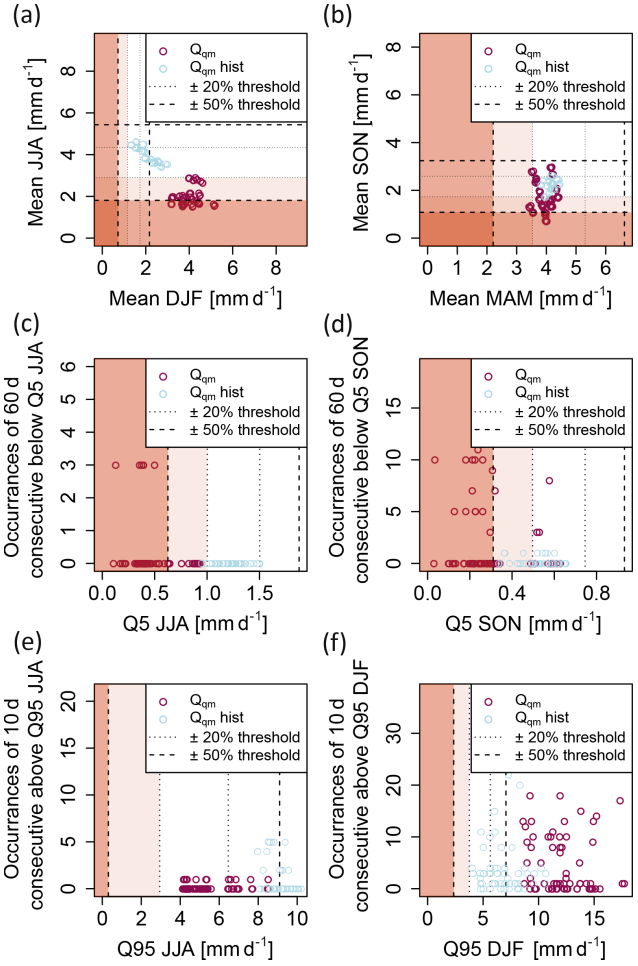



**Figure 8.** Situations leading to the greatest stress on Groupe E's operations are depicted by the comparisons of low-flow, high-flow, and seasonal-flow indices for the Vernex catchment. Two time periods are compared in the left and right columns: **(a–c)** 1980–2009 and **(d–f)** 2070–2099. Panels **(d–f)** show simulations under the influence of RCP8.5. Panels **(a)** and **(b)** depict seasonal flows: mean winter flow (DJF) versus mean summer flow (JJA). Panels **(c)** and **(d)** depict low flows: Q5 summer flows (JJA) versus the occurrence of consecutive days below Q5. Panels **(e)** and **(f)** depict high flows: Q95 winter flows (DJF) versus the occurrence of consecutive days above Q95. For all plots, two thresholds, which were provided by Groupe E, are included: ±20 % and ±50 %. Shading from white to progressively darker red tones indicates the lowest (white) to highest (dark red) levels of stress placed on Groupe E's operations based on the relationship between the indices.

2013; Savelsberg et al., 2018). This motivated the selection of the electricity demand indices for this study. Although a temperature-based electricity demand approach is inherently limited, it is a sensible way to initiate the discussion on future changes in climate, the electricity market, and electricity consumption behaviors. Our analysis using temperature indices suggests that electricity demand in summer and fall may increase (Fig. 7b), which will be difficult to satisfy using only inflow entering the reservoirs (as it is expected to decrease over these seasons), implying that ownership in other electricity sectors may be needed to respond to future electricity demand. The Swiss Energy Strategy 2050 stipulates that the deficit left from the decommissioning of nuclear power plants should be partially compensated for by an increase in hydropower production. However, as Switzerland has almost reached its maximum capacity for hydropower production, renewables (e.g., wind and photovoltaics) are expected to play a significant role in supplementing the deficit left by the phaseout of nuclear power (Redondo and Van Vliet, 2015).

Storage hydropower plants have the ability to release water and generate energy in response to electricity prices in order to create revenue (Savelsberg et al., 2018). A flexible operation mode could allow Groupe E to capitalize on peak prices, as electricity prices are expected to become more volatile due to the increased contribution of renewable energy sources to the electricity market (Anghileri et al., 2018). However, regulations regarding water rights in some countries limit the ability of hydropower operations to change their mode of operation (e.g., the water rights would have to be renegotiated to enable the plant operators to update the design of their installation; Gaudard et al., 2016). More flexibility (e.g., the duration of the contract, the installation design and capacity, and low-flow requirements) could be incorporated into the water concession, as the vested rights within a concession cannot currently undergo important changes once agreed upon. The flexibility of concessions is discussed by Gaudard (2015), who argues that concessions should last 40 years rather than 80 years; the abovementioned study also points out that the more flexible the concession, the more it gains in value.

### 5.2 Benefits of developing tailored projections by following a stakeholder-centered approach

Involving stakeholders in the modeling and figure design provided key benefits and insights (Addor et al., 2015). It revealed, for instance, that the indices chosen by impact modelers are not necessarily well suited to support decision-making. Although standard indices, such as the long-term mean monthly distribution of inflow, are useful, given the complexity of the concession renegotiation process, a single index or non-tailored indices are of limited use. Instead, indices need to be chosen to bridge the gap between the global-scale climate change phenomenon and concerns and vulnerabilities at the regional to local level. This, for instance, led to the selection of a less common index – consecutive days of low flows – which enabled us to explore a critical vulnerability of hydropower operations that is often overlooked by top-down impact studies. The importance of tailored projections is especially apparent when compared to the existing literature on climate change impacts on hydropower production in Switzerland. The expected mean monthly inflow changes for the Vernex and Montsalvens catchments are most

comparable to projections for the nearby Emme catchment simulated by Addor et al. (2014). However, given the local-scale information needed for hydropower management and concession negotiation, indices beyond the long-term mean monthly cycle are needed. Finger et al. (2012) produced hydrological projections for the Saas Fee region in Switzerland, but these are not directly useable by Groupe E, as the hydrological indices they analyzed are not specific enough for concession negotiations nor is the alpine region they cover expected to respond in the same way to climate change as region of Groupe E's catchments.

Groupe E managers expressed that our collaboration enabled them to envision the impacts of climate change at the local level and to prepare for the impacts they may experience as electricity managers. Groupe E is interested in similar studies for other catchments, and they are considering an investment in additional hydrological projections in the future. They stressed the importance of having access to inflow projections in order to begin the process of climate change adaptation and to prepare for critical conversations prior to official negotiations. They stated that, compared with the generic information they have access to, this collaboration made the climate change phenomenon more real and that the figures we co-produced provided them with a clear picture of the likely impacts of climate change on their activities. This highlights the benefits of the direct inclusion of stakeholders to anticipate and efficiently prepare for future climate change impacts.

### 5.3 Visualizing climate change impact projections and their uncertainties to inform decision-making

Characterizing and visualizing projections uncertainty played a central role during this project, as hydropower managers must negotiate their water concessions despite an abundance of uncertainty (Gaudard et al., 2016). At the onset of the project, we made sure to understand how Groupe E interprets the uncertainty intervals associated with the inflow projections. Uncertainties associated with model calibration (parametric uncertainty) and multi-model ensembles (structural uncertainty) were already familiar to Groupe E, because they routinely utilize an ensemble of streamflow forecasts and account for these uncertainties in their day-to-day operations. Groupe E explained that they consciously consider the width of uncertainty bands compared to the mean change in order to assess the robustness of changes. For instance, Fig. 5a shows the magnitude of Q5 over the JJA and SON seasons between the historical (1980–2009) and future (2070–2099) periods. The spread of the projections is reflected by the width of the boxplots. Figures 5a shows a clear change between historical and future low flows, where all future ensemble members exceed the $-50\,\%$ threshold specified by Groupe E. This result represents a profit loss for Groupe E because there will likely be less water available for turbination, and, if turbinated, it will be at a lower efficiency. In other cases, when results are less definitive, Groupe E stated that the mean (or median) of the projections is most useful to them.

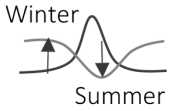
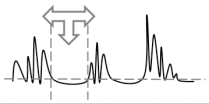
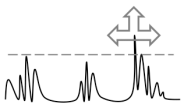
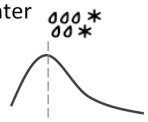
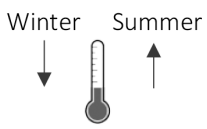
Our visuals were subject to multiple rounds of feedback, where different variables were compared and shown to Groupe E so that we were able to tell a meaningful story. For instance, a decision-analytic summary figure was created based on Fig. 2 in Brown et al. (2012) and was initially proposed to Groupe E. This type of figure uses two axes to show changes in two selected variables and indicates which decision is optimal for different regions of this two-dimensional space. Groupe E pointed out that, given their situation, the value of this type of visual is limited as it is too simple to display the numerous considerations influencing the concession renewal. Instead, Fig. 5 in Broderick et al. (2019) was used as a basis for Fig. 8 to succinctly visualize changes in a series of key indices in relation to the specified thresholds. A summary table of the main opportunities and adaptation options was also provided to Groupe E (Table 4). In addition, given this project's focus on hydrological changes relevant for hydropower operations, we selected climate models based on historical hydrological performance. Some climate models were found to generate unrealistic simulations of discharge or snow processes and were not used for further analysis (see Table 3). Models that produced unrealistic snow processes were excluded given that the cold biases associated with the unbridled snow accumulation may impact the climate change signal of the surrounding grid cells and, thus, provide unreliable projections of hydrological change.

### 5.4 Limitations and next steps

Concession negotiations have many facets and although hydrological changes are important, they only partially determine the profitability of hydropower operations. This study focused on hydroclimatic changes using a range of streamflow indices. We did not account for the uncertainties related to the development of the European or Swiss electricity market. Instead, we used a simple method to estimate future electricity demand solely based on air temperature. Nevertheless, this study points out that despite the uncertainties involved, quantifying the supply of future water resources and providing an estimate of changes to demand (based on changes to air temperature) improves the information currently available to electricity managers and is useful for their concession negotiations.

There is now a need to complement this analysis with a more economical analysis, focused on the future electricity demand and on the evolution of the electricity market. A collaboration between climate impact and energy–economic modeling (e.g., Anghileri et al., 2018; Savelsberg et al., 2018) seems to be the natural next step. Economical studies often aggregate all climate change impacts by focusing on the profitability of the reservoir and only consider changes in the seasonal cycle. In contrast, this study shows

**Table 4.** The major opportunities and risks for hydropower operations presented in relation to the hydrological and climatological considerations for concession renewal.

| | | | | | |
|---|---|---|---|---|---|
| Vulnerability | Overall decrease in annual inflows and seasonal change in the distribution of inflow | Increase in the duration of low flows while less water is carried by low flows | Seasonal change in the behavior of high flows | Meltwater mixing with rain events while reservoir levels approach annual maximum | Electricity demand – decrease over winter and increase over summer |
| Index | • Seasonal means | • Q5<br>• Consecutive days below Q5 | • Q95<br>• Consecutive days above Q95 | • Rain versus snow contribution to runoff | • $H_{DD}$<br>• $C_{DD}$ |
| Change in index | Increase in long-term mean monthly inflow over winter<br><br>Decrease in long-term mean monhtly inflow over summer and fall<br>  | Reduction of the magnitude of water carried by Q5 by 50% over summer and fall<br><br>Duration of low flows below Q5 will likely extend as long as 80–90 consecutive days<br>  | Duration of high flows above Q95 will likely extend for as long as 20 consecutive days over winter<br><br>Likely decrease in summer high flows<br>  | Peak annual contribution from snowpack will likely shift from May to April<br><br>Rain will likely increase its contribution to inflow during the winter<br>  | Likely decrease in demand for electricity over winter<br><br>Likely increase in demand for electricity over summer and fall<br>  |
| Opportunity – adaptation option | <u>Winter</u><br>• Increase in inflows when electricity prices have historically been high could result in profit<br><br><u>Summer</u><br>• Summer inflows more unreliable<br>• Reduced inflows will likely make it harder to meet reservoir-level requirements | <u>Summer</u><br>• Less water & longer periods of low flows will likely make it harder to meet residual flow requirements<br>• Negotiate new/flexible residual flow requirements as part of concession<br>• These projections could provide a basis for price reduction for water fee | <u>Winter</u><br>• Increase in inflows when prices are historically high could result in profit<br><br><u>Summer</u><br>• High flows are likely to be unreliable to offset extended periods of low flows<br>• Diversify electricity mix | <u>Spring</u><br>• Fast runoff when reservoir levels are high could make reservoir level management precarious<br><br><u>All seasons</u><br>• High-intensity events could be turned to profit if forecasted early – lock in trade/sale based on forecast & bolster forecast systems | <u>All seasons</u><br>• Future electricity demand is highly uncertain<br>• Suggest flexible concession terms<br>• A potential increase in dependency on renewables likely means intermittency in production, which could represent a challenge in meeting future demand |

how linking stakeholder vulnerabilities to changes in individual indices offers an approachable means to evaluate adaptation measures compared with a lumped profit/loss figure. New research projects would benefit from involving a wider range of stakeholders. A collaboration between hydrologists, economists, and stakeholders, such as cantonal authorities, environmental interest groups, hydropower operations specialists, and electricity market traders, would help to support concession negotiations and foster the sustainable development of hydropower.

Additional streamflow indices would be useful to Groupe E, in particular those related to the magnitude and duration of flooding. Future work should include rare and potentially damaging flooding events. The indices and thresh-

olds chosen by Groupe E should not be assumed to be adequate for all hydropower climate change adaptation studies. Instead, we advocate for stakeholder involvement early in future studies so that indices, modeling chains, and results can be tailored for decision-making. Finally, future work could also involve the characterization of sources of uncertainty not considered in this study, such as hydrological model uncertainty and natural variability.

## 6 Conclusions

This study demonstrates the benefits of involving stakeholders early in climate change impact studies. While most hydroclimatic impact studies explore streamflow changes in isolation and rarely address their implications for water management (Gaudard et al., 2013; Hänggi and Weingartner, 2012), this project went beyond a usual top-down analysis and addressed the specific needs and concerns of stakeholders. We worked with representatives from a hydroelectricity company, and we asked them to describe their main vulnerabilities to hydroclimatic variations; together, we then selected hydrological and electricity demand indices to characterize future impacts. These results enabled us to identify likely key challenges and opportunities for hydropower operations under climate change and to provide guidance on the upcoming water concession negotiations. Our projections indicate a significant increase in inflow over the winter period when electricity prices have historically been at their highest. In contrast, a reduction in summer inflows is expected and will represent a challenge, given the possible increase in electricity demand for cooling as a result of higher temperature. Our projections of low flows provide a basis to support the negotiation of new residual flow requirements. The projected increase in high flows over the winter period could represent an opportunity if this water can be captured and turbinated at optimal times or at prearranged prices. The involvement of stakeholders early on in the project was vital to ensuring that the results and figures of this study were directly useful for their concession negotiations and provide insights into how their operations are likely to be impacted by climate change.

This study is timely as many electricity managers are currently faced with renegotiating their water concessions in the context of climate change and an uncertain electricity market. However, studies such as Tonka (2015) note, there has been a "striking lack of attention paid to climate change impacts on water resources availability in relicensure procedures". We show that although many uncertainties exist, given the multi-decade length of a concession, it is crucial for climate change to be considered at the onset of concession negotiations. The analysis presented here is transferable to other water management entities and provides guidance for other climate change projects that strive to follow a stakeholder-centered approach and deliver projections useful for decision-making.

*Code and data availability.* EURO-CORDEX data can be accessed via different European data nodes and are available at https://www.hzg.de/ms/euro-cordex/060378/index.php.en (Earth System Grid Federation, 2019). The TabsD and RhiresD interpolated products used to create the historical reference climate time series are available from MeteoSwiss. The streamflow time series representing the inflow to the Lac de Montsalvens and Lac du Vernex was provided by Groupe E.

*Supplement.* The Supplement for this article includes the projections of the Montsalvens reservoir and the additional time periods analyzed (the time periods from 2020 to 2049, from 2045 to 2074, and from 2070 to 2099 were analyzed; however, given length considerations, only projections covering 2070–2099 are included within the paper). In addition, the following materials are available online for those who wish to carry out their own hydrological climate change impact analysis: (i) a download of HBV hydrological model (https://www.geo.uzh.ch/en/units/h2k/Services/HBV-Model.html, last access: May 2020) and (ii) an encyclopedia chapter that introduces the steps of a hydrological climate change impact assessment (Hakala et al., 2020). The supplement related to this article is available online at: https://doi.org/10.5194/hess-24-3815-2020-supplement.

*Author contributions.* KH and NA designed this study based on previous exchanges between NA and Groupe E. KH refined the scope of the project with Groupe E over the course of several meetings. KH performed the climate change impact analysis, with oversight from NA and JS. Writing of the paper was led by KH with feedback from all coauthors; finalization of the paper primarily occurred between KH and NA.

*Competing interests.* The authors declare that they have no conflict of interest.

*Acknowledgements.* The authors thank Marc Vis for support related to the HBV model. We also thank Hannes Weigt for providing useful insights into the Swiss water fee framework. We acknowledge the Federal Office of Meteorology and Climatology MeteoSwiss and the European Coordinated Regional Downscaling Experiment (EURO-CORDEX) for the atmospheric and climate modeled data, respectively. Valuable comments by the editor and the reviewers helped to improve this paper.

*Financial support.* This research has been supported by the Swiss National Science Foundation (grant nos. 200020_156606 and P400P2_180791).

*Review statement.* This paper was edited by Markus Hrachowitz and reviewed by two anonymous referees.

## References


Addor, N., Rössler, O., Köplin, N., Huss, M., Weingartner, R., and Seibert, J.: Robust changes and sources of uncertainty in the projected hydrological regimes of Swiss catchments, Water Resour. Res., 50, 1–22, https://doi.org/10.1002/2014WR015549, 2014.

Addor, N., Ewen, T., Johnson, L., Çöltekin, A., Derungs, C., and Muccione, V.: From products to processes: Academic events to foster interdisciplinary and iterative dialogue in a changing climate, Earth's Futur., 3, 289–297, https://doi.org/10.1002/2015EF000303, 2015.

Anghileri, D., Castelletti, A., and Burlando, P.: Alpine Hydropower in the Decline of the Nuclear Era: Trade-Off between Revenue and Production in the Swiss Alps, J. Water Resour. Plan. Manag., 144, 1–11, https://doi.org/10.1061/(ASCE)WR.1943-5452.0000944, 2018.

Barry, M., Baur, P., Gaudard, L., Giuliani, G., Hediger, W., Romerio, F., Schillinger, M., Schumann, R., Voegeli, G., and Weigt, H.: The Future of Swiss Hydropower A Review on Drivers and Uncertainties, SCCER CREST Work. 3 Energy Policy, Mark. Regul., 1–49, 2015.

Bergström, S.: Development and Application of a Conceptual Runoff Model for Scandinavian Catchments, SMHI, Norrköping, 1976.

Betz, R., Geissmann, T., Kosch, M., Schillinger, M., and Weigt, H.: The Design of Variable Water Fees and its Impact on Swiss Hydropower Companies and Resource Owners, Work Packag. 3 Energy Policy, Mark. Regul. SCCER CREST, 1–21, 2019.

Bitterli, T., George, M., Matousek, F., Christe, R., Brändli, R., Frey, D., and Tripet, J.-P.: Groundwater Resources, Hydrol. Atlas Switzerland, Plate 8.6, 6–9, 2004.

Brigode, P., Oudin, L., and Perrin, C.: Hydrological model parameter instability: A source of additional uncertainty in estimating the hydrological impacts of climate change?, J. Hydrol., 476, 410–425, https://doi.org/10.1016/j.jhydrol.2012.11.012, 2013.

Broderick, C., Murphy, C., Wilby, R. L., Matthew, T., Prudhomme, C., and Adamson, M.: Using a Scenario – Neutral Framework to Avoid Potential Maladaptation to Future Flood Risk Water Resources Research, Water Resour. Res., 55, 1–26, https://doi.org/10.1029/2018WR023623, 2019.

Brown, C., Ghile, Y., Laverty, M., and Li, K.: Decision Scaling?: Linking bottom-up vulnerability analysis with climate projections in the water sector, Water Resour. Res., 48, 1–12, https://doi.org/10.1029/2011WR011212, 2012, 2012.

Coron, L., Andréassian, V., Perrin, C., Lerat, J., Vaze, J., Bourqui, M. and Hendrickx, F.: Crash testing hydrological models in contrasted climate conditions: An experiment on 216 Australian catchments, Water Resour. Res., 48, 1–17, https://doi.org/10.1029/2011WR011721, 2012.

Earth System Grid Federation: project name: CORDEX, available at: https://www.hzg.de/ms/euro-cordex/060378/index.php.en, last access: May 2019.

Etter, S., Addor, N., Huss, M., and Finger, D.: Climate change impacts on future snow, ice and rain runoff in a Swiss mountain catchment using multi-dataset calibration, J. Hydrol. Reg. Stud., 13, 222–239, https://doi.org/10.1016/j.ejrh.2017.08.005, 2017.

Fan, Y.: Are catchments leaky?, WIRES Water, 6, e1386, https://doi.org/10.1002/wat2.1386, 2019.

Finger, D., Heinrich, G., Gobiet, A., and Bauder, A.: Projections of future water resources and their uncertainty in a glacierized catchment in the Swiss Alps and the subsequent effects on hydropower production during the 21st century, Water Resour. Res., 48, 1–20, https://doi.org/10.1029/2011WR010733, 2012.

FOEN: Effects of Climate Change on Water Resources and Waters, Synth. Rep. "Climate Chang. Hydrol. Switzerland" Proj., Federal Office for the Environment (FOEN), Bern, Switzerland, 2012.

Frei, C.: Interpolation of temperature in a mountainous region using nonlinear profiles and non-Euclidean distances, Int. J. Climatol., 34, 1585–1605, https://doi.org/10.1002/joc.3786, 2014.

Frei, C. and Schär, C.: A precipitation climatology of the Alps from high-resolution rain-gauge observations, Int. J. Climatol., 18, 873–900, https://doi.org/10.1002/(SICI)1097-0088(19980630)18:8<873::AID-JOC255>3.0.CO;2-9, 1998.

Frei, P., Kotlarski, S., Liniger, M. A., and Schär, C.: Future snowfall in the Alps: projections based on the EURO-CORDEX regional climate models, The Cryosphere, 12, 1–24, https://doi.org/10.5194/tc-12-1-2018, 2018.

Gaudard, L.: Pumped-storage project: A short to long term investment analysis including climate change, Renew. Sustain. Energy Rev., 49, 91–99, https://doi.org/10.1016/j.rser.2015.04.052, 2015.

Gaudard, L., Gilli, M., and Romerio, F.: Climate Change Impacts on Hydropower Management, Water Resour. Manag., 27, 5143–5156, https://doi.org/10.1007/s11269-013-0458-1, 2013.

Gaudard, L., Gabbi, J., Bauder, A., and Romerio, F.: Long-term Uncertainty of Hydropower Revenue Due to Climate Change and Electricity Prices, Water Resour. Manag., 30, 1325–1343, https://doi.org/10.1007/s11269-015-1216-3, 2016.

Gaudard, L., Voegeli, G., and Romerio, F.: Hydropower Investment Profitability under Different Water Fee Systems, Swiss National Science Foundation, Geneva, 2018a.

Gaudard, L., Avanzi, F., and De Michele, C.: Seasonal aspects of the energy-water nexus: The case of a run-of-the-river hydropower plant, Appl. Energ., 210, 604–612, https://doi.org/10.1016/j.apenergy.2017.02.003, 2018b.

Gudmundsson, L.: qmap: Statistical transformations for post-processing climate model output, R package version 1.0-4, 2016.

Gudmundsson, L., Bremnes, J. B., Haugen, J. E., and Engen-Skaugen, T.: Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations – a comparison of methods, Hydrol. Earth Syst. Sci., 16, 3383–3390, https://doi.org/10.5194/hess-16-3383-2012, 2012.

Gupta, H. V, Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, J. Hydrol., 377, 80–91, https://doi.org/10.1016/j.jhydrol.2009.08.003, 2009.

Hakala, K., Addor, N., and Seibert, J.: Hydrological modeling to evaluate climate model simulations and their bias correction, J. Hydrometeorol., 19, 1321–1337, https://doi.org/10.1175/JHM-D-17-0189.1, 2018.

Hakala, K., Addor, N., Teutschbein, C., Vis, M., Dakhlaoui, H., and Seibert, J.: Hydrological Modeling of Climate Change Impacts, in: Encyclopedia of Water, edited by: Maurice, P., John Wiley and Sons, Inc., https://doi.org/10.1002/9781119300762.wsts0062, 2020.

Hänggi, P. and Weingartner, R.: Variations in Discharge Volumes for Hydropower Generation in Switzerland, Water Resour. Manag., 26, 1231–1252, https://doi.org/10.1007/s11269-011-9956-1, 2012.

Kienast, F., Huber, N., Hergert, R., Bolliger, J., Moran, L. S., and Hersperger, A. M.: Conflicts between decentralized renewable electricity production and landscape services – A spatially-explicit quantitative assessment for Switzerland, Renew. Sustain. Energy Rev., 67, 397–407, https://doi.org/10.1016/j.rser.2016.09.045, 2017.

Klemeš, V.: Operational testing of hydrological simulation models, Hydrol. Sci. J., 31, 13–24, https://doi.org/10.1080/02626668609491024, 1986.

Köplin, N., Schädler, B., Viviroli, D., and Weingartner, R.: Seasonality and magnitude of floods in Switzerland under future climate change, Hydrol. Process., 28, 2567–2578, https://doi.org/10.1002/hyp.9757, 2014.

Lindström, G., Johansson, B., Persson, M., Gardelin, M., and Bergström, S.: Development and test of the distributed HBV-96 hydrological model, J. Hydrol., 201, 272–288, https://doi.org/10.1016/S0022-1694(97)00041-3, 1997.

Lopez, A., Fung, F., New, M., Watts, G., Weston, A., and Wilby, R. L.: From climate model ensembles to climate change impacts and adaptation: A case study of water resource management in the southwest of England, Water Resour. Res., 45, 1–21, https://doi.org/10.1029/2008WR007499, 2009.

Mauch, C. and Reynard, E.: The Evolution of National Water Regimes in Switzerland, Evol. Natl. Water Regimes Eur. Environ. Policy, 40, 293–328, https://doi.org/10.1007/978-1-4020-2484-9_9, 2004.

Meinshausen, M., Smith, S. J., Calvin, K., Daniel, J. S., Kainuma, M. L. T., Lamarque, J.-F., Matsumoto, K., Montzka, S. A., Raper, S. C. B., Riahi, K., Thomson, A., Velders, G. J. M., and van Vuuren, D. P. P.: The RCP greenhouse gas concentrations and their extensions from 1765 to 2300, Clim. Change, 109, 213–241, https://doi.org/10.1007/s10584-011-0156-z, 2011.

Menzel, L., Lang, H., and Rohmann, M.: Mean Annual Actual Evaporation, Hydrol. Atlas Switzerland, Plate 4.1, 3–4, 1999.

Nash, J. E. and Sutcliffe, J. V: River flow forecasting through conceptual models part I – A discussion of principles, J. Hydrol., 10, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

Redondo, P. D. and Van Vliet, O.: Modelling the energy future of Switzerland after the phase out of nuclear power plants, Energ. Proc., 76, 49–58, https://doi.org/10.1016/j.egypro.2015.07.843, 2015.

Savelsberg, J., Schillinger, M., Schlecht, I., and Weigt, H.: The impact of climate change on Swiss hydropower, Sustainability, 10, 1–23, https://doi.org/10.3390/su10072541, 2018.

Schaefli, B.: Projecting hydropower production under future climates: a guide for decision-makers and modelers to interpret and design climate change impact assessments, Wiley Interdiscip. Rev. Water, 2, 271–289, https://doi.org/10.1002/wat2.1083, 2015.

Schaefli, B., Hingray, B., and Musy, A.: Climate change and hydropower production in the Swiss Alps: quantification of potential impacts and related modelling uncertainties, Hydrol. Earth Syst. Sci., 11, 1191–1205, https://doi.org/10.5194/hess-11-1191-2007, 2007.

Schwarb, M.: The alpine precipitation climate: Evaluation of a high-resolution analysis scheme using comprehensive rain-gauge data, Swiss Federal Institute of Technology Zurich, Zurich, 2000.

Seibert, J.: Multi-criteria calibration of a conceptual runoff model using a genetic algorithm, Hydrol. Earth Syst. Sci., 4, 215–224, https://doi.org/10.5194/hess-4-215-2000, 2000.

Seibert, J. and Vis, M. J. P.: Teaching hydrological modeling with a user-friendly catchment-runoff-model software package, Hydrol. Earth Syst. Sci., 16, 3315–3325, https://doi.org/10.5194/hess-16-3315-2012, 2012.

Smiatek, G., Kaspar, S., and Kunstmann, H.: Hydrological Climate Change Impact Analysis for the Figeh Spring near Damascus, Syria, J. Hydrometeorol., 14, 577–593, https://doi.org/10.1175/jhm-d-12-065.1, 2012.

SWV: Heimfall und Neukonzessionierung von Wasserkraftwerken, SWV, Baden, 2012.

Tonka, L.: Hydropower license renewal and environmental protection policies: a comparison between Switzerland and the USA, Reg. Environ. Chang., 15, 539–548, https://doi.org/10.1007/s10113-014-0598-8, 2015.

Vano, J. A., Scott, M. J., Voisin, N., Stöckle, C. O., Hamlet, A. F., Mickelson, K. E. B., Elsner, M. M. G., and Lettenmaier, D. P.: Climate change impacts on water management and irrigated agriculture in the Yakima River Basin, Washington, USA, Clim. Change, 102, 287–317, https://doi.org/10.1007/s10584-010-9856-z, 2010.

Vano, J. A., Arnold, J. R., Nijssen, B., Clark, M. P., Wood, A. W., Gutmann, E. D., Addor, N., Hamman, J., and Lehner, F.: DOs and DON'Ts for using climate change information for water resource planning and management: guidelines for study design, Clim. Serv., 12, 1–13, https://doi.org/10.1016/j.cliser.2018.07.002, 2018.

Vicuna, S. and Dracup, J. A.: The evolution of climate change impact studies on hydrology and water resources in California, Clim. Change, 82, 327–350, https://doi.org/10.1007/s10584-006-9207-2, 2007.

Wilby, R. L. and Dessai, S.: Robust adaptation to climate change, Weather, 65, 180–185, https://doi.org/10.1002/wea.504, 2010.

Zierl, B. and Bugmann, H.: Global change impacts on hydrological processes in Alpine catchments, Water Resour. Res., 41, 1–13, https://doi.org/10.1029/2004WR003447, 2005.

Zubler, E. M., Fischer, A. M., Fröb, F., and Liniger, M. A.: Climate change signals of CMIP5 general circulation models over the Alps – impact of model selection, Int. J. Climatol., 36, 3088–3104, https://doi.org/10.1002/joc.4538, 2016.

# Acknowledgements

I would like to thank the many people who have contributed to the content of this dissertation through meetings, coffee breaks, and office discussions. This thesis would not have been possible without the generosity of my colleagues' time, humor, patience, and distraction.

To begin, I would like to thank Jan Seibert, my supervisor. Its hard to believe that that our scientific collaboration started all the way back in Colorado. Thank you for the opportunity to come to Switzerland to conduct a PhD and for your generous support and encouragement to attend workshops and conferences throughout Europe. I couldn't have asked for a better experience here in your group - it's truly been a pleasure to work with you.

I owe a deep debt of gratitude to Nans Addor, my second supervisor. I am incredibly lucky to have had the opportunity to work with you and equally to have gotten to know you through our many chats. I will always be thankful for the many lessons that I've learned and for your tireless encouragement.

Thanks to Reinhard Furrer for the valuable input and nice discussions during the committee meetings.

Thanks to Marc Vis for being a go-to source of help for trouble shooting throughout the different dissertation projects.

Thanks also to Sven Kotlarski for his availability and for sharing his expertise on climate models.

I truly enjoyed my time spent with my colleagues/friends from H2K. I'm grateful for how close we all became through the laughs, practical jokes on Marc, horrible Christmas gift exchanges, and great discussions.

Thank you to my family, and in particular my mother, for all the skype chats and sharing in the excitement of my life abroad.

And lastly, Rick - thank you for all the in-between moments of my PhD - the office band, the meerkat visits, our strolls around Irchel Park. Thank you for your calming presence, clever mind, curiosity, and all the great/horrible jokes that you force me to listen to. Ik houd van jou.

This thesis received financial support from the Swiss National Science Foundation under the grant no. 200020_156606.

Cover: Current (green) and future (red) long-term mean monthly streamflow cycles are compared. Projections of future streamflow are under the influence of climate change. Cover kindly painted by Linda Assendelft.